

GenCore version 5.1.4\_P5\_4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 3, 2003, 21:30:10 ; Search time 4788 Seconds  
(without alignments)  
10606.599 Million cell updates/sec

Title: US-10-017-621-3

Perfect score: 1745

Sequence: 1 tggagcgcgcglaaaggatg.....gttcacgtgccactgttcc 1745

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 841850

Minimum DB seq length: 0  
Maximum DB seq length: 50

Post-processing: Minimum Watch 0%  
Maximum Watch 100%  
Listing first 45 summaries

Database :

GenEmbl:  
1: gb\_ba:\*  
2: gb\_hgt:\*  
3: gb\_in:\*  
4: gb\_om:\*  
5: gb\_ov:\*  
6: gb\_pat:\*  
7: gb\_ph:\*  
8: gb\_pl:\*  
9: gb\_pr:\*  
10: gb\_ro:\*  
11: gb\_sts:\*  
12: gb\_sy:\*  
13: gb\_un:\*  
14: gb\_vl:\*  
15: em\_ba:\*  
16: em\_fun:\*  
17: em\_hum:\*  
18: em\_in:\*  
19: em\_mu:\*  
20: em\_om:\*  
21: em\_or:\*  
22: em\_ov:\*  
23: em\_pat:\*  
24: em\_ph:\*  
25: em\_pl:\*  
26: em\_ro:\*  
27: em\_sts:\*  
28: em\_un:\*  
29: em\_vl:\*  
30: em\_htg\_hum:\*  
31: em\_htg\_inv:\*  
32: em\_htg\_other:\*  
33: em\_htg\_mus:\*  
34: em\_htg\_pln:\*  
35: em\_htg\_rod:\*  
36: em\_htg\_man:\*  
37: em\_htg\_vrt:\*  
38: em\_sy:\*  
39: em\_htgo\_hum:\*  
40: em\_htgo\_mus:\*  
41: em\_htgo\_other:\*

Score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	23.8	1.4	49	9	HSHCAK
2	22.8	1.3	50	6	A93721
3	22.4	1.3	50	6	A93722
4	22.2	1.3	47	6	I84671
5	22	1.3	50	6	AX159452
6	21.6	1.2	31	6	AX248673
7	21.4	1.2	42	6	AX182243
8	21.4	1.2	42	6	AX182243
9	21.4	1.2	46	6	AX320208
10	21.4	1.2	46	6	AX320208
11	21.4	1.2	46	6	AX320208
12	21.4	1.2	46	6	AX320208
13	21	1.2	31	6	AX248015
14	21	1.2	46	6	AX186238
15	20.8	1.2	46	6	AX186238
16	20.6	1.2	21	6	AX153998
17	20.6	1.2	45	6	AX153998
18	20.6	1.2	45	6	AX153998
19	20.6	1.2	45	6	AX153998
20	20.4	1.2	45	6	AX153998
21	20.4	1.2	45	6	AX153998
22	20.4	1.2	45	6	AX153998
23	20.2	1.2	40	6	AX153998
24	20.2	1.2	40	6	AX153998
25	20.2	1.2	40	6	AX153998
26	20.2	1.2	40	6	AX153998
27	20.2	1.2	40	6	AX153998
28	20.2	1.2	40	6	AX153998
29	20.2	1.2	40	6	AX153998
30	20	1.1	36	6	AX153998
31	20	1.1	36	6	AX153998
32	20	1.1	36	6	AX153998
33	20	1.1	36	6	AX153998
34	20	1.1	36	6	AX153998
35	19.8	1.1	46	6	AX036348
36	19.8	1.1	46	6	AX036348
37	19.8	1.1	46	6	AX036348
38	19.8	1.1	46	6	AX036348
39	19.6	1.1	48	23	BD004595
40	19.6	1.1	42	6	AX153233
41	19.6	1.1	42	6	AX153233
42	19.6	1.1	45	6	AX153233
43	19.6	1.1	45	6	AX153233
44	19.6	1.1	50	6	AX032970
45	19.6	1.1	50	6	AX032970

# ALIGNMENTS

RESULT 1  
HSHCAK  
LOCUS HSHCAK  
DEFINITION H. sapiens mRNA for Cdk activating kinase.  
ACCESSION X76171  
VERSION X76171.1  
KEYWORDS activating kinase; protein kinase.  
SOURCE Homo sapiens  
ORGANISM Homo sapiens  
REFERENCE 1 (bases 1 to 49)  
AUTHORS Hall, F. L.  
TITLE Direct Submission  
JOURNAL Submitted (08-NOV-1993) F.L. Hall, Childrens Hospital Los Angeles, PRI 08-AUG-1995

Pred. No. is the number of results predicted by chance to have a

50 CAGAACCTCCACCTCCAGAACCTCCACCTCCAGAACCTCCACCTCCG 1

RESULT 3	LOCUS	50 bp	DNA	linear	PAT 27-JAN-2000
A93722	A93722	Sequence 9 from Patent WO9732891.			

ACCESSION	A93722
VERSION	A93722.1
KEYWORDS	GI:6741894
SOURCE	.
ORGANISM	unidentified.
	unidentified
	unclassified
REFERENCE	1 (bases 1 to 50)
AUTHORS	Rose-John,S.
TITLE	CONJUGATE FOR MODIFYING INTERACTIONS BETWEEN PROTEINS
JOURNAL	Patent: WO 9732891-A 9 12-SEP-1997; ANGSWANDT GENTECHNOLOGIE SYST (DE); ROSE JOHN STEFAN (DE)

```

REFERENCE 1 (bases 1 to 50)
unclassified.

```

AUTHORS	Rose-John, S.
TITLE	CONJUGATE FOR MODIFYING INTERACTIONS BETWEEN PROTEINS
JOURNAL	PATENT: WO 9732891-A 12-SEP-1997;
FEATURES	ANGEMANDTE GENTECNOLOGIE SYST (DE); ROSE JOHN STEFAN
	Location/Qualifiers
source	1..50

BASE COUNT	13 a	26 c	4 g	7 t
ORIGIN				

Query Match	1.3%	Score 22.4	DB 5	Length 50
Best Local Similarity	66.7%	Pred. No. 1.6e+06		
Matches	32	Conservative 0	Mismatches 16	Indels
958	CGCGAGAGGGTGTACACTCAGACCTCAAGCCCGAGAACCTGCTCATC	1005		
DB	2	CGACGAAACCTCCACCTCCGAGACCTCCACCTCCGAGAACCTCCACCTC	49	

LOCUS	SEQUENCE	BP	DNA	LINEAR	PAT
184671	Sequence 5	from patent	US 5696086		
184671	Accession				
184671.1	GI:3022191				
184671.1	Keywords				
184671.1	Source				
184671.1	Organism				

REFERENCE 1 (bases 1 to 47)

TITLE	Methods and kits using macrophage stimulating protein
JOURNAL	Patent: US 5696086-A 5 09-DEC-1997;
FEATURES	Location/Qualifiers
SOURCE	J. . 47

BASE COUNT	8 a	16 c	/organism
100	100	100	100
200	200	200	200
300	300	300	300
400	400	400	400
500	500	500	500
600	600	600	600
700	700	700	700
800	800	800	800
900	900	900	900
1000	1000	1000	1000

[illegible]

2 GGACGAAFLCACCATGGGGTTC

<b>RESULT 5</b>			
<b>AXI59452</b>			
<b>XLOCUS</b>			
<b>DEFINITION</b>	AXI59452	50 bp	DNA
<b>DEFINITION</b>	Sequence 2780 from Patent WO0140521.		linear PAT 22-JUN-2001

DEFINITION Sequence 2780 from Patent WO0140521.  
ACCESSION AX159452  
VERSION AX159452.1 GI:14540783  
KEYWORDS

SOURCE	ORGANISM
human.	
human.	Homo sapiens
Eukaryota.	Metazoa.
Eukaryota.	Chordata.
Eukaryota.	Vertebrata.
Eukaryota.	Euteleostomi.

1

REFERENCE 1 (bases 1 to 50)  
 AUTHORS Shimketa, R.A. and Leach, M.  
 TITLE Nucleic acids containing single nucleotide polymorphisms and methods of use thereof  
 JOURNAL Patent: WO 0140521-A 2780 07-JUN-2001;  
 FEATURES Location/Qualifiers  
 source 1..50  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 misc\_feature 25..26  
 /note="Nucleotide deleted between bases 25 and 26  
 Accession number c942160243"  
 misc\_feature 26  
 /note="2 of 2 allelic variants (2779 is other entry)"  
 BASE COUNT 7 a 22 c 13 g 8 t  
 ORIGIN

Query Match 1.38; Score 22; DB 6; Length 50;  
 Best Local Similarity 73.74; Pred. No. 2e+06;  
 Matches 28; Conservative 0; Mismatches 10; Indels 0; Gaps 0;  
 QY 1642 CGCTGAGGATGACACACCCCTACAGGGCAGCC 1679  
 Db 11 CTGCTGAGCCCTGACACACCCCTCTGCGGCCCC 48

RESULT 6  
 AX248673  
 LOCUS AX248673 31 bp DNA linear PAT 28-SEP-2001  
 DEFINITION Sequence 752 from Patent WO0166800.  
 ACCESSION AX248673  
 VERSION AX248673.1 GI:15863296  
 KEYWORDS  
 SOURCE human.  
 ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 31)  
 AUTHORS Cargill, M., Ireland, J.S. and Lander, E.S.  
 TITLE Human single nucleotide polymorphisms  
 JOURNAL Patent: WO 0166800-A 752 13-SEP-2001;  
 WHITEHEAD INSTITUTE FOR BIOMEDICAL RESEARCH (US)  
 FEATURES Location/Qualifiers  
 source 1..31  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"

BASE COUNT 8 a 11 c 7 g 4 t 1 others  
 ORIGIN

Query Match 1.24; Score 21.6; DB 6; Length 31;  
 Best Local Similarity 80.04; Pred. No. 2.6e+06;  
 Matches 24; Conservative 1; Mismatches 5; Indels 0; Gaps 0;  
 QY 979 GACCTCAGCCCCAGACCTGCTCATTCAAC 1008  
 Db 2 GACATCAGCCCCAGACCTGCTGTTGGAC 31

RESULT 7  
 AX182243/C  
 LOCUS AX182243 42 bp DNA linear PAT 06-AUG-2001  
 DEFINITION Sequence 53 from Patent WO0142441.  
 ACCESSION AX182243  
 VERSION AX182243.1 GI:15133518  
 KEYWORDS  
 SOURCE synthetic construct.  
 ORGANISM artificial sequences.  
 REFERENCE 1 (bases 1 to 42)  
 AUTHORS Reddy, S.I., Sadhu, L.I., Shukla, V.C. and Ferraiolo, G.I.  
 TITLE Plasmid transformation

JOURNAL Patent: WO 0142441-A 53 14-JUN-2001;  
 International Centre for Genetic Engineering and Biotechnology (IT)  
 FEATURES Location/Qualifiers  
 source 1..42  
 /organism="synthetic construct"  
 /db\_xref="taxon:32630"  
 /note="Primer"  
 BASE COUNT 15 a 10 c 11 g 6 t  
 ORIGIN

Query Match 1.24; Score 21.4; DB 6; Length 42;  
 Best Local Similarity 80.64; Pred. No. 2.8e+06;  
 Matches 25; Conservative 0; Mismatches 6; Indels 0; Gaps 0;  
 QY 270 ACGTCTCTCTCTGGGAACTTCTTCTGCA 300  
 Db 35 ACGTAGGGTCTCTGGCGACCTTCGATCTGCA 5

RESULT 8  
 AX382049/C  
 LOCUS AX382049 42 bp DNA linear PAT 18-MAR-2002  
 DEFINITION Sequence 53 from Patent WO0206497.  
 ACCESSION AX382049  
 VERSION AX382049.1 GI:19576871  
 KEYWORDS  
 SOURCE synthetic construct.  
 ORGANISM synthetic construct.  
 REFERENCE 1  
 AUTHORS Reddy, V.S. and Sadhu, L.  
 TITLE Transplasmatic plants  
 JOURNAL Patent: WO 0206497-A 53 24-JAN-2002;  
 International Centre for Genetic Engineering and Biotechnology (IT)  
 FEATURES Location/Qualifiers  
 source 1..42  
 /organism="synthetic construct"  
 /db\_xref="taxon:32630"  
 /note="Primer"

BASE COUNT 15 a 10 c 11 g 6 t  
 ORIGIN

Query Match 1.24; Score 21.4; DB 6; Length 42;  
 Best Local Similarity 80.64; Pred. No. 2.8e+06;  
 Matches 25; Conservative 0; Mismatches 6; Indels 0; Gaps 0;  
 QY 270 ACGTCTCTCTCTGGGAACTTCTTCTGCA 300  
 Db 35 ACGTAGGGTCTCTGGCGACCTTCGATCTGCA 5

RESULT 9  
 AR032544  
 LOCUS AR032544 46 bp DNA linear PAT 29-SEP-1999  
 DEFINITION Sequence 156 from patent US 5869241.  
 ACCESSION AR032544  
 VERSION AR032544.1 GI:5948149  
 KEYWORDS  
 SOURCE Unknown.  
 ORGANISM Unclassified.

REFERENCE 1 (bases 1 to 46)  
 AUTHORS Edwards, C.A., Cantor, C.R., Andrews, B.M., Turin, L.M. and Fry, K.E.  
 TITLE Method of determining DNA sequence preference of a DNA-binding molecule  
 JOURNAL Patent: US 5869241-A 156 09-FEB-1999;  
 FEATURES Location/Qualifiers  
 source 1..46  
 /organism="unknown"  
 BASE COUNT 9 a 14 c 16 g 7 t  
 ORIGIN

Query Match 1.24; Score 21.4; DB 6; Length 46;





```

/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT      7 a      13 c      18 g      6 t      2 others
ORIGIN

Query Match      1.28; Score 21; DB 6; Length 46;
Best Local Similarity 71.13; Pred. NO. 3.5e+06;
Matches 27; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

OY 550 AAGCCCTCAGCGCGCTCCGTCGTGTCAGCCTATC 587
      |||| ||| || ||||| ||| || |||||
DB 41 AAGCCTCTCTGCAGCGCCGCCGCCGAGTCTCCTATC 4

RESULT 15
A98791/c
LOCUS      A98791
DEFINITION Sequence 24 from Patent WO9910358.
ACCESSION  A98791
VERSION    A98791.1 GI:6781812
KEYWORDS   .
SOURCE     unidentified.
ORGANISM   unidentified.
REFERENCE  1 (bases 1 to 46)
AUTHORS   Hegemann P.
TITLE     METHOD FOR PRODUCING NUCLEIC ACID POLYMERS
JOURNAL   Patent: WO 9910358-A 24 04-MAR-1999;
          HEGEMANN PETER (DE)
FEATURES   Location/Qualifiers
            source
              1..46
              /organism="unidentified"
              /db_xref="taxon:32644"
BASE COUNT      9 a      13 c      16 g      8 t
ORIGIN

Query Match      1.28; Score 20.8; DB 6;
Best Local Similarity 78.18; Pred. NO. 3.9e+06;
Matches 25; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

OY 500 TGCGTCAGCGGCTACTCGAGCAGCTGACCTC 531
      |||| ||||| ||| || ||| ||| ||
DB 37 TGCCGAGCGGCTACTCGAGGAGCGGCCATC 6

Search completed: March 4, 2003, 00:06:28
Job time : 4794 secs

```

GenCore version 5.1.4.p5\_4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 3, 2003, 22:34:05 ; Search time 2628 seconds  
(without alignments)  
10753.853 Million cell updates/sec

Title: US-10-017-621-3

Perfect score: 1745

Sequence: 1 tggaaagcagcgttaagcgtg.....gttcactgccactgtgcc 1745

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 102860

Minimum DB seq length: 0

Maximum DB seq length: 50

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:\*

1: em\_estba:\*

2: em\_esthum:\*

3: em\_estin:\*

4: em\_estmu:\*

5: em\_estov:\*

6: em\_estpl:\*

7: em\_estro:\*

8: em\_esti:\*

9: gb\_est1:\*

10: gb\_est2:\*

11: gb\_hic:\*

12: gb\_est3:\*

13: gb\_est4:\*

14: gb\_est5:\*

15: em\_estfun:\*

16: em\_estom:\*

17: gb\_gss:\*

18: em\_gss\_hum:\*

19: em\_gss\_inv:\*

20: em\_gss\_pin:\*

21: em\_gss\_vrt:\*

22: em\_gss\_fun:\*

23: em\_gss\_mam:\*

24: em\_gss\_mus:\*

25: em\_gss\_other:\*

26: em\_gss\_pro:\*

27: em\_gss\_rod:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	45	2.6	46	14	N78054
2	28	1.6	28	14	R38968
3	23.2	1.3	48	10	AW247978
4	21.4	1.2	50	9	AW107934
5	21.2	1.2	16	17	R2146286
6	21	1.2	50	9	AW102877

c	7	21	1.2	50	9	AW105237
c	8	21	1.2	50	13	BM397711
	9	20.8	1.2	46	17	A2993993
	10	20.6	1.2	48	17	A2311362
c	11	20.6	1.2	50	9	AW106960
c	12	20.4	1.2	47	17	A2311536
c	13	20.2	1.2	45	17	A2985975
c	14	20.2	1.2	49	9	AW204601
c	15	20	1.1	49	9	AA1475974
	16	20	1.1	50	17	BH811451
	17	19.8	1.1	49	14	W39000
c	18	19.4	1.1	50	9	AW104829
	19	19.2	1.1	44	12	BG422154
	20	19.2	1.1	49	17	A2450961
c	21	19.2	1.1	50	9	AW103357
c	22	19.2	1.1	50	9	AW103358
c	23	19.2	1.1	50	9	AW103359
c	24	19.2	1.1	50	9	AW103361
	25	19.2	1.1	50	9	AW103381
	26	19.2	1.1	50	9	AW103915
	27	19.2	1.1	50	9	AW106349
c	28	19.2	1.1	50	14	T74703
	29	19	1.1	43	9	AI591257
	30	19	1.1	49	17	A2423762
	31	19	1.1	50	9	AW107320
c	32	18.8	1.1	34	9	AA972479
c	33	18.8	1.1	43	17	BH857724
	34	18.8	1.1	46	17	BH790015
	35	18.8	1.1	50	9	AW102939
	36	18.8	1.1	50	9	AW103583
c	37	18.8	1.1	50	9	AW104587
c	38	18.8	1.1	50	9	AW105918
c	39	18.6	1.1	44	17	TA165G050
c	40	18.6	1.1	45	9	AI250043
	41	18.6	1.1	49	17	A2966392
	42	18.6	1.1	50	9	AW103382
	43	18.6	1.1	50	9	AW103553
c	44	18.6	1.1	50	9	AW104162
c	45	18.6	1.1	50	13	BI910989

#### ALIGNMENTS

RESULT 1  
N78054  
LOCUS  
DEFINITION  
YV71905.r1 Soares fetal liver spleen INFLS Homo sapiens CDNA clone  
IMAGE:248216 5', similar to gb:X66363 SERINE/THREONINE-PROTEIN  
KINASE PCTAIRE-1 (HUMAN);, mRNA sequence.

ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM

N78054  
N78054.1  
GI:1240755

human

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

AUTHORS

1 (bases 1 to 46)  
Hillier, L., Lennon, G., Becker, M., Bonaldo, M.P., Chiapelli, B.,  
Chissoe, S., Dietrich, N., DuBuque, T., Favello, A., Gish, W., Hawkins,  
M., Hultman, M., Kucaba, T., Lacy, M., Le, M., Le, N., Mardis, E., Moore,  
B., Morris, M., Parsons, J., Prange, C., Rifkin, L., Rohlfing, T., Moore,  
Schellenberg, K., Soares, M.B., Tan, F., Thierry-Mieg, J., Trevaskis, E.,  
Underwood, K., Woldmann, P., Waterston, R., Wilson, R. and Marra, M.  
Generation and analysis of 280,000 human expressed sequence tags  
Genome Res. 6 (9), 807-828 (1996)

TITLE

JOURNAL

MEDLINE

COMMENT

Contact: Wilson RK

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

Tel: 314 286 1800

Fax: 314 286 1810

Email: est@watson.wustl.edu

This clone is available royalty-free through LLNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information. Trace considered overall poor quality. Insert Length: 1438 Std Error: 0.00 Seq primer: reverse ET High quality sequence stop: 1.

## FEATURES

Location/Qualifiers  
1. .46  
/organism="Homo sapiens"  
/db\_xref="GDB:3797462"  
/db\_xref="taxon:9606"  
/clone="IMAGE:248216"  
/clone\_lib="Soares fetal liver spleen INFLS"  
/sex="male"  
/dev\_stage="20 week-post conception fetus"  
/lab\_host="DH10B (ampicillin resistant)"  
/note="Organ: Liver and Spleen; Vector: pT7T3D (Pharmacia) with a modified polylinker; Site\_1: Pac I; Site\_2: Eco RI; 1st strand cDNA was primed with a Pac I - oligo(dT) primer [5' AACCTGGAGAAATTAATTAAGATCTTTTCTTTTCTTTT 3'], double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Pac I and cloned into the Pac I and Eco RI sites of the modified pT7T3 vector. Library went through one round of normalization. Library constructed by Bento Soares and M.Fatima Bonaldo."

## BASE COUNT

6 a 20 c 7 g 12 t  
ORIGIN

Query Match 2.6%; Score 45; DB 14; Length 46;  
Best Local Similarity 97.8%; Pred. No. 0.78;  
Matches 45; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1700 ACTCTGCTGCTACCTGCTGAGCCATGTTACCTGCCACTGTGCC 1745  
|||||  
Db 1 ACTCTGCTGCTACCTGCTGAGCCATGTTACCTGCCACTGTGCC 46

## RESULT 2

R38968/c  
LOCUS  
DEFINITION  
IMAGE:25073 3' similar to gb:X66363 SERINE/THREONINE-PROTEIN KINASE  
PCTAIRE-1 (HUMAN); mRNA sequence.

ACCESSION  
R38968  
VERSION  
R38968.1 GI:796424  
KEYWORDS  
EST.  
SOURCE  
Homo sapiens

ORGANISM  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 28)

REFERENCE  
AUTHORS  
Hillier, L., Clark, M., Dubuque, T., Elliston, K., Hawkins, M., Holman, R., Hultman, M., Kucaba, T., Le, N., Lennon, G., Marra, M., Parsons, J., Rifkin, L., Rohlfing, T., Soares, M., Tan, F., Trevisan, E., Waterston, R., Williamson, A., Woldmann, P., and Wilson, R.  
The Marsh-Merck EST Project  
Unpublished (1995)  
Contact: Wilton RK  
Washington University School of Medicine  
444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: est@wustl.edu

## TITLE

JOURNAL  
COMMENT  
High quality sequence starts: 1 High quality sequence stops: 1  
Source: IMAGE Consortium, LLNL This clone is available royalty-free through LLNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information. Trace considered overall poor quality  
Insert Length: 1349 Std Error: 0.00  
Seq primer: -21m13  
High quality sequence stop: 1.  
Location/Qualifiers  
1. .48

## FEATURES

source

/organism="Homo sapiens"  
/db\_xref="GDB:397420"  
/db\_xref="taxon:9606"  
/clone="IMAGE:25073"  
/clone\_lib="Soares infant brain IN1B"  
/sex="female"  
/dev\_stage="73 days post natal"  
/lab\_host="DH10B (ampicillin resistant)"  
/note="Organ: whole brain; Vector: Lefmid BA; Site\_1: Not I; Site\_2: Hind III; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5' AACCTGGAGAAATGCGCCGCGAGATTTTCTTTTCTTTT 3'], double-stranded cDNA was ligated to Hind III adaptors (Pharmacia), digested with Not I and directionally cloned into the Not I and Hind III sites of the Lefmid BA vector. Library went through one round of normalization. Library constructed by Bento Soares and M.Fatima Bonaldo."

BASE COUNT  
8 a 5 c 11 g 4 t  
ORIGIN

Query Match 1.6%; Score 28; DB 14; Length 28;  
Best Local Similarity 100.0%; Pred. No. 9.1e+03;  
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1716 CCGAGCCATGTTACCTGCCACTGTCT 1743  
|||||  
Db 28 CCGAGCCATGTTACCTGCCACTGTCT 1

RESULT 3  
LOCUS  
DEFINITION  
IMAGE:25073 5prime NIH\_MGC\_7 Homo sapiens cDNA clone IMAGE:2819657 5', mRNA sequence.

ACCESSION  
R38968  
VERSION  
R38968.1 GI:5591066  
KEYWORDS  
EST.  
SOURCE  
Homo sapiens

ORGANISM  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 48)

REFERENCE  
AUTHORS  
NIH-MGC http://mgi.nci.nih.gov/  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Other ESTs: 2819657.3prime  
Contact: Robert Strausberg, Ph.D.  
Email: cgapbs@mail.nih.gov  
Tissue Procurement: DCTD/DMP cDNA Library Preparation: Ling Hong/Rubin Laboratory cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Berkeley MGC sequencing project Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www.bio.llnl.gov/bbrp/image/image.html Base Calling / Quality Trimming: cross\_match from University of Washington Genome Center PHRAP suite. Poly-T Identification: patmatch.pl from Berkeley Drosophila Genome Project. University of Washington Genome Center: http://www.genome.washington.edu Low Quality Sequence: 8 contiguous PHRED high quality bases following vector sequence. Very Low Quality Sequence: Trace file contained 48 contiguous distinct peaks following vector sequence.  
Plate: LLCM2 row: C column: 18  
High quality sequence stop: 8.  
Location/Qualifiers  
1. 48

FEATURES  
source

/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:2819657"  
/clone\_lib="NIH\_MGC\_7"  
/tissue\_type="small cell carcinoma"  
/cell\_line="MGC3"  
/lab\_host="DH10B (phage-resistant)"



```

ACCESSION   AU102877
VERSION     AU102877.1  GI:13552398
SOURCE      EST.
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 50)
AUTHORS     Suzuki,Y., Taira,H., Tsunoda,T., Mizushima-Sugano,J., Sese,J., Hata
            ,H., Ota,T., Isogai,T., Tanaka,T., Morishita,S., Okubo,K., Sakaki
            ,Y., Nakamura,Y., Suyama,A. and Sugano,S.
TITLE       Diverse transcriptional initiation revealed by fine, large-scale
            mapping of mRNA start sites
JOURNAL     EMBO Rep. 2 (5), 388-393 (2001)
MEDLINE     21270072
COMMENT     Contact: Yutaka Suzuki
            Department of Virology
            Institute of Medical Science, University of Tokyo
            4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
            Email: yusuzuki@ims.u-tokyo.ac.jp
            Suzuki,Y., Yoshitomo-Nakagawa,K., Maruyama,K., Suyama,A. and Sugano
            ,S. Construction and characterization of a full length-enriched and
            a 5'-end-enriched cDNA library. Gene 200 (1-2), 149-156 (1997).

FEATURES             source
     source           1..50
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="COL03073"
                     /clone_lib="Sugano Homo sapiens cDNA library"
                     /note="Differential display comparison of untreated and
                     dimethylformate treated U937 cells"

BASE COUNT      11 a 20 c 10 g
ORIGIN
Query Match      1.2%  Score 21;  DB 9;  Length 50;
Best Local Similarity 82.8%;  Pred. No. 6.3e+05;
Matches 24;  Conservative 0;  Mismatches 5;  Indels 0;  Gaps 0;

QY  491 ACATCCGCTGCTGAGGCTACCTGGAG 519
      ||||| ||||| ||||| ||||| ||
Db    1 ACATCCGCTGCTGAGACCTTCGCGAG 29

RESULT 7
AU105237/c
LOCUS         AU105237 Sugano Homo sapiens cDNA library Homo sapiens cDNA clone
DEFINITION   HRC08919, mRNA sequence.
ACCESSION   AU105237
VERSION     AU105237.1  GI:13554758
KEYWORDS     EST.
SOURCE      human.
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 50)
AUTHORS     Suzuki,Y., Taira,H., Tsunoda,T., Mizushima-Sugano,J., Sese,J., Hata
            ,H., Ota,T., Isogai,T., Tanaka,T., Morishita,S., Okubo,K., Sakaki
            ,Y., Nakamura,Y., Suyama,A. and Sugano,S.
TITLE       Diverse transcriptional initiation revealed by fine, large-scale
            mapping of mRNA start sites
JOURNAL     EMBO Rep. 2 (5), 388-393 (2001)
MEDLINE     21270072
COMMENT     Contact: Yutaka Suzuki
            Department of Virology
            Institute of Medical Science, University of Tokyo
            4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
            Email: yusuzuki@ims.u-tokyo.ac.jp
            Suzuki,Y., Yoshitomo-Nakagawa,K., Maruyama,K., Suyama,A. and Sugano
            ,S. Construction and characterization of a full length-enriched and
            a 5'-end-enriched cDNA library. Gene 200 (1-2), 149-156 (1997).

FEATURES             source
     source           1..50
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="COL03073"
                     /clone_lib="Sugano Homo sapiens cDNA library"
                     /note="Differential display comparison of untreated and
                     dimethylformate treated U937 cells"

```

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="HRC08919"
/clone_lib="Sugano Homo sapiens cDNA library"
/note="Differential display comparison of untreated and
dimethylformate treated U937 cells"

BASE COUNT      13 a 10 c 22 g 5 t
ORIGIN
Query Match      1.2%  Score 21;  DB 9;  Length 50;
Best Local Similarity 82.8%;  Pred. No. 6.3e+05;
Matches 24;  Conservative 0;  Mismatches 5;  Indels 0;  Gaps 0;

QY  558 CAGCCGCCCTCGCTGCTGTCAGCCTAT 586
      ||||| ||||| ||||| ||||| |||||
Db    35 CAGCTCCGCCGCCCATCTCTGCGCCTAT 7

RESULT 8
BM397711/c
LOCUS         BM397711 50 bp mRNA linear EST 17-JAN-2002
DEFINITION   Tetrahymena thermophila cDNA, mRNA sequence.
ACCESSION   BM397711
VERSION     BM397711.1  GI:18197764
KEYWORDS     EST.
SOURCE      Tetrahymena thermophila.
            Tetrahymena thermophila
            Eukaryota; Alveolata; Ciliophora; Oligohymenophorea;
            Hymenostomatida; Tetrahymenina; Tetrahymena.
REFERENCE   1 (bases 1 to 50)
AUTHORS     Turkewitz,A.P., Karier,K.M., Jahn,C., Orlas,E., Kirk,K.E., Frankel
            ,J. and Klobutcher,L.
            EST from Tetrahymena thermophila, strain CU428.1, growing cells
            Unpublished (2002)
            Contact: Turkewitz AP
            Molecular Genetics and Cell Biology
            University of Chicago
            920 E. 58th Street, Chicago, IL 60637, USA
            Tel: 773 702 4374
            Fax: 773 702 3172
            Email: apturkew@midway.uchicago.edu
            Seq primer: T3.

FEATURES             source
     source           1..50
                     /organism="Tetrahymena thermophila"
                     /strain="CU428.1"
                     /db_xref="taxon:5911"
                     /clone_lib="Chilcoat/Turkewitz cDNA (large fraction)"
                     /note="Vector: Bluescript2 SK+, Details on library
                     preparation can be found in Chilcoat and Turkewitz (2001)
                     Proc. Natl. Acad. Sci USA, 98: 8709-8713."

BASE COUNT      10 a 7 c 23 g 2 t 8 others
ORIGIN
Query Match      1.2%  Score 21;  DB 13;  Length 50;
Best Local Similarity 60.0%;  Pred. No. 6.3e+05;
Matches 27;  Conservative 0;  Mismatches 18;  Indels 0;  Gaps 0;

QY  82 CCCCAGGCTGTGAGTTGTCTCGCGCGCCCGCCGCGATGCGCATG 126
      ||||| ||||| ||||| ||||| ||||| |||||
Db    50 CCCCAGGACTCCAGCTTTGTCCCNCCNNNNNNGCTCTCCCTG 6

RESULT 9
AZ993993
LOCUS         AZ993993 46 bp DNA linear GSS 27-APR-2001
DEFINITION   Clone UUC2M240279E13 P, DNA sequence.
ACCESSION   AZ993993
VERSION     AZ993993.1  GI:13865220
KEYWORDS     GSS.
SOURCE      house mouse.

```











GenCore version 5.1.4.p5.4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 3, 2003, 22:39:05 : Search time 133 Seconds  
(without alignments)  
8182.066 Million cell updates/sec

Title: US-10-017-621-3  
Perfect score: 1745  
Sequence: 1 tgggaagcagcgttaaggatg.....gttcaactgccacttgctc 1745

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Capext 1.0

Searched: 460893 seqs, 311809382 residues

Total number of hits satisfying chosen parameters: 254638

Minimum DB seq length: 0  
Maximum DB seq length: 50

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published\_Applications\_NA:

1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq.\*  
2: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq.\*  
3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq.\*  
4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq.\*  
5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq.\*  
6: /cgn2\_6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq.\*  
7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq.\*  
8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq.\*  
9: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq.\*  
10: /cgn2\_6/ptodata/2/pubpna/US09\_PUBCOMB.seq.\*  
11: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq.\*  
12: /cgn2\_6/ptodata/2/pubpna/US10\_PUBCOMB.seq.\*  
13: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq.\*  
14: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	21.8	1.2	48	9 US-10-034-444-6	Sequence 6, Appl
2	21.6	1.2	31	10 US-09-801-274-752	Sequence 752, App
3	21.4	1.2	45	9 US-10-029-413A-25	Sequence 25, Appl
4	21	1.2	31	10 US-09-801-274-94	Sequence 94, Appl
5	20.4	1.2	48	9 US-09-943-722-47	Sequence 47, Appl
6	20	1.1	45	10 US-09-147-142-11	Sequence 11, Appl
7	20	1.1	45	10 US-09-147-142-12	Sequence 12, Appl
8	19.8	1.1	40	10 US-09-263-959-758	Sequence 758, App
9	19.6	1.1	42	10 US-09-790-417-235	Sequence 235, App
10	19.4	1.1	45	12 US-10-073-256-35	Sequence 27, Appl
11	19.4	1.1	45	12 US-10-073-256-33	Sequence 35, Appl
12	19.2	1.1	31	10 US-09-801-274-517	Sequence 517, App
13	19.2	1.1	46	10 US-09-263-959-121	Sequence 121, App
14	19	1.1	43	9 US-09-376-940-23	Sequence 23, Appl
15	19	1.1	45	10 US-09-818-066-32	Sequence 32, Appl
16	19	1.1	47	9 US-10-118-231-9	Sequence 9, Appl
17	18.8	1.1	48	9 US-09-840-277-104	Sequence 104, App
18	18.8	1.1	48	9 US-09-840-277-105	Sequence 105, App
19	18.8	1.1	48	10 US-09-753-436-67	Sequence 67, Appl

Sequence 107, App  
Sequence 15295, A  
Sequence 3433, Ap  
Sequence 5256, Ap  
Sequence 18, Appl  
Sequence 31, Appl  
Sequence 31, Appl  
Sequence 9, Appl  
Sequence 233, App  
Sequence 4, Appl  
Sequence 8, Appl  
Sequence 10, Appl  
Sequence 14, Appl  
Sequence 16, Appl  
Sequence 3226, Ap  
Sequence 2, Appl  
Sequence 2, Appl  
Sequence 25, Appl  
Sequence 24, Appl  
Sequence 25, Appl  
Sequence 33, Appl  
Sequence 34, Appl  
Sequence 3267, Ap  
Sequence 3311, Ap  
Sequence 3491, Ap

49 9 US-09-840-277-107  
25 10 US-09-866-108-15295  
23 18.6 1.1 48 9 US-09-864-785-3433  
22 18.4 1.1 41 9 US-09-938-842A-5256  
24 18.4 1.1 48 10 US-09-761-534A-18  
25 18.2 1.0 36 9 US-10-219-248-31  
26 18.2 1.0 36 9 US-10-219-247-31  
27 18.2 1.0 36 10 US-09-855-723-31  
28 18.2 1.0 50 9 US-09-943-722-9  
29 18 1.0 42 10 US-09-790-417-233  
30 18 1.0 43 12 US-10-043-142-4  
31 18 1.0 46 9 US-10-026-914-8  
32 18 1.0 46 9 US-10-026-914-10  
33 18 1.0 46 9 US-10-026-914-16  
34 18 1.0 48 9 US-09-864-785-3226  
35 18 1.0 35 9 US-10-051-989-2  
36 17.8 1.0 35 9 US-09-861-097-2  
37 17.8 1.0 42 10 US-09-865-807-25  
38 17.8 1.0 45 12 US-10-073-256-24  
39 17.8 1.0 45 12 US-10-073-256-25  
40 17.8 1.0 45 12 US-10-073-256-33  
41 17.8 1.0 45 12 US-10-073-256-34  
42 17.8 1.0 48 9 US-09-864-785-3267  
43 17.8 1.0 48 9 US-09-864-785-3311  
44 17.8 1.0 48 9 US-09-864-785-3491  
45 17.8 1.0 48 9 US-09-864-785-3491

## ALIGNMENTS

## RESULT 1

US-10-034-444-6  
: Sequence 6, Application US/10054444  
: Patent No. US2002016432A1

GENERAL INFORMATION:  
: APPLICANT: Guyre, Paul M.

: APPLICANT: Goldstein, Joel

: APPLICANT: Wu, Zining

: APPLICANT: Sun, Wanwen

: TITLE OF INVENTION: Recombinant Cat Allergen, Fel d1, Expressed In

: TITLE OF INVENTION: Baculovirus for Diagnosis and Treatment of Cat Allergy

: FILE REFERENCE: DC-0118

: CURRENT APPLICATION NUMBER: US/10/054,444

: PRIOR FILING DATE: 2002-01-22

: PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/410,963

: NUMBER OF SEQ ID NOS: 6

: SOFTWARE: Patent Ver. 2.0

: SEQ ID NO 5

: LENGTH: 48

: TYPE: DNA

: ORGANISM: Artificial Sequence

: FEATURE:

: OTHER INFORMATION: Description of Artificial Sequence: Synthetic

US-10-034-444-6

Query Match 1.2% Score 21.8; DB 9;  
Best Local Similarity 70.7%; Pred. No. 9.5e+03;  
Matches 29; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

OY 962 AGAAGTCTACACGACCTCAAGCCCGACAGCTCTC 1002

DB 2 AGAAGTCTCTCCAGACCTCTCTCCACGAGACCTCTC 42

## RESULT 2

US-09-801-274-752

: Sequence 752, Application US/09801274

: Patent No. US20020032319A1

: GENERAL INFORMATION:

: APPLICANT: Cargill, Michele

: APPLICANT: Ireland, James S.



TOPOLOGY: LINEAR  
US-09-943-722-47

Query Match 1.18; Score 20.4; DB 9; Length 48;  
Best Local Similarity 71.18; Pred. No. 2.3e+04;  
Matches 27; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

OY 325 GAGATTGTCACGAGGACTTGAAGATGGGTCTGATGG 362  
||||| ||||| ||||| ||||| ||||| ||||| |||||  
DB 7 GAGACGGTGGCCGGAAGTTGAAGCCGGGATGATGG 44

## RESULT 6

US-09-147-142-11  
; Sequence 11 Application US/09147142  
; Patent No. US20020018749A1  
; GENERAL INFORMATION:  
; APPLICANT: HUDSON, Peter John  
; APPLICANT: KORTT, Alex Andrew  
; APPLICANT: IRVING, Robert Alexander  
; APPLICANT: ATWELL, John Leslie  
; TITLE OF INVENTION: HIGH AVIDITY POLYVALENT AND POLYSPECIFIC REAGENTS  
; FILE REFERENCE: 016786/0212  
; CURRENT APPLICATION NUMBER: US/09/147,142  
; CURRENT FILING DATE: 1999-03-05  
; EARLIER APPLICATION NUMBER: PCT/AU98/00212  
; EARLIER FILING DATE: 1998-03-26  
; EARLIER APPLICATION NUMBER: AU PO 5917  
; EARLIER FILING DATE: 1997-03-27  
; NUMBER OF SEQ ID NOS: 32  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 11  
; LENGTH: 45  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: synthetic  
; OTHER INFORMATION: oligonucleotide used to insert codon between VH  
; OTHER INFORMATION: and VL domains of NC10 scfv-0  
US-09-147-142-11

Query Match 1.18; Score 20; DB 10; Length 45;  
Best Local Similarity 65.98; Pred. No. 2.8e+04;  
Matches 29; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

OY 643 GGTACCTATGCCACCGTCTACAAAGGCAAGCAAGCTCACAGA 686  
||||| ||||| ||||| ||||| ||||| ||||| |||||  
DB 1 GGGACACCGTACCGCTCTCCGGTGGTATCGAGCTCACACA 44

## RESULT 7

US-09-147-142-12/c  
; Sequence 12 Application US/09147142  
; Patent No. US20020018749A1  
; GENERAL INFORMATION:  
; APPLICANT: HUDSON, Peter John  
; APPLICANT: KORTT, Alex Andrew  
; APPLICANT: IRVING, Robert Alexander  
; APPLICANT: ATWELL, John Leslie  
; TITLE OF INVENTION: HIGH AVIDITY POLYVALENT AND POLYSPECIFIC REAGENTS  
; FILE REFERENCE: 016786/0212  
; CURRENT APPLICATION NUMBER: US/09/147,142  
; CURRENT FILING DATE: 1999-03-05  
; EARLIER APPLICATION NUMBER: PCT/AU98/00212  
; EARLIER FILING DATE: 1998-03-26  
; EARLIER APPLICATION NUMBER: AU PO 5917  
; EARLIER FILING DATE: 1997-03-27  
; NUMBER OF SEQ ID NOS: 32  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 12  
; LENGTH: 45  
; TYPE: DNA  
; ORGANISM: Artificial Sequence

FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: synthetic  
; OTHER INFORMATION: oligonucleotide used to insert codon between VH  
; OTHER INFORMATION: and VL domains of NC10 scfv-0  
US-09-147-142-12

Query Match 1.18; Score 20; DB 10; Length 45;  
Best Local Similarity 65.98; Pred. No. 2.8e+04;  
Matches 29; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

OY 643 GGTACCTATGCCACCGTCTACAAAGGCAAGCAAGCTCACAGA 686  
||||| ||||| ||||| ||||| ||||| ||||| |||||  
DB 45 GGGACACCGTACCGCTCTCCGGTGGTATCGAGCTCACACA 2

## RESULT 8

US-09-263-959-758  
; Sequence 758 Application US/091631959  
; Patent No. US20020150891A1  
; GENERAL INFORMATION:  
; APPLICANT: Hood, Leroy E.  
; APPLICANT: ROVEN, Lee  
; APPLICANT: KOOP, Ben F.  
; TITLE OF INVENTION: DIAGNOSTIC AND THERAPEUTIC COMPOSITIONS AND METHODS WHICH U  
; NUMBER OF SEQUENCES: 1279  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Seed and Berry LLP  
; STREET: 6300 Columbia Center, 701 Fifth Avenue  
; CITY: Seattle  
; STATE: Washington  
; COUNTRY: US  
; ZIP: 98104-7092  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA: US/09/263,959  
; FILING DATE: 05-MAR-1999  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: McHastere, David D.  
; REGISTRATION NUMBER: 33,963  
; REFERENCE/DOCKET NUMBER: 920010.426C2  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (206) 622-4900  
; TELEFAX: (206) 682-8031  
; INFORMATION FOR SEQ ID NO: 758:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 40 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
US-09-263-959-758

Query Match 1.18; Score 19.8; DB 10; Length 40;  
Best Local Similarity 77.48; Pred. No. 3e+04;  
Matches 24; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

OY 1689 CTTCCTCTGCTTACTCTCTGCTACCTGCTGCTG 1719  
||||| ||||| ||||| ||||| ||||| ||||| |||||  
DB 3 CTTCCTCTCTCTCTCTCTCTCTCTCTCTGCTGCTG 33

## RESULT 9

US-09-790-417-235/c  
; Sequence 235 Application US/09790417  
; Patent No. US20010031470A1  
; GENERAL INFORMATION:  
; APPLICANT: Shultz, John W  
; APPLICANT: Lewis, Martin K.  
; APPLICANT: Lieppe, Donna

```
; APPLICANT: Mandrekar, Michelle
; APPLICANT: Kephart, Daniel
; APPLICANT: Rhodes, Richard B.
; APPLICANT: Andrews, Christine A.
; APPLICANT: Hartnett, James R.
; APPLICANT: Gu, Trent
; APPLICANT: Olson, Ryan J.
; APPLICANT: Wood, Keith W.
; APPLICANT: Welch, Roy
; TITLE OF INVENTION: Nucleic Acid Detection
; FILE REFERENCE: Pro-103 6868/75528
; CURRENT APPLICATION NUMBER: US/09/790,417
; CURRENT FILING DATE: 2001-02-22
; PRIOR APPLICATION NUMBER: 09/358,972
; PRIOR FILING DATE: 1999-07-21
; PRIOR APPLICATION NUMBER: 09/042,287
; PRIOR FILING DATE: 1998-03-13
; NUMBER OF SEQ ID NOS: 290
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 235
; LENGTH: 42
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:probe for oligo
; OTHER INFORMATION: 54
US-09-790-417-235

Query Match      1.1%  Score 19.6; DB 10; Length 42;
Best Local Similarity 66.7%; Pred. No. 3.5e+04;
Matches 28; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY 759 GTCCCTGCTCAAGGACCTCAACACGCGCAACATCGTTACGCT 800
   |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 42 CTACCTGCTAAATGAACCTCACCCACGATATCATCAACGACT 1

RESULT 10
US-10-073-256-27/c
; Sequence 27, Application US/10073256
; Patent No. US20020120408A1
; GENERAL INFORMATION:
; APPLICANT: Kreiswirth, Barry N
; APPLICANT: Nadich, Steven M
; TITLE OF INVENTION: System and Method for Tracking and Controlling Infections
; FILE REFERENCE: 19124.0002
; CURRENT APPLICATION NUMBER: US/10/073,256
; CURRENT FILING DATE: 2002-02-13
; NUMBER OF SEQ ID NOS: 80
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 27
; LENGTH: 45
; TYPE: DNA
; ORGANISM: Enterococcus faecalis
US-10-073-256-27

Query Match      1.1%  Score 19.4; DB 10; Length 45;
Best Local Similarity 64.4%; Pred. No. 4.1e+04;
Matches 29; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 1288 ATCTGTCACAGGAGGTTCAGACATACAACTACCCCAAGTAC 1332
   |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 45 AACCAGGAAACCACTAGTACACGCGGAAAGCAGCACTCAAGTAC 1

RESULT 11
US-10-073-256-35/c
; Sequence 35, Application US/10073256
; Patent No. US20020120408A1
; GENERAL INFORMATION:
; APPLICANT: Kreiswirth, Barry N
; APPLICANT: Nadich, Steven M
; TITLE OF INVENTION: System and Method for Tracking and Controlling Infections
```

```
; FILE REFERENCE: 19124.0002
; CURRENT APPLICATION NUMBER: US/10/073,256
; CURRENT FILING DATE: 2002-02-13
; NUMBER OF SEQ ID NOS: 80
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 35
; LENGTH: 45
; TYPE: DNA
; ORGANISM: Enterococcus faecalis
US-10-073-256-35

Query Match      1.1%  Score 19.4; DB 12; Length 45;
Best Local Similarity 64.4%; Pred. No. 4.1e+04;
Matches 29; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 1288 ATCTGTCACAGGAGGTTCAGACATACAACTACCCCAAGTAC 1332
   |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 45 AACCAGTGAACCACTAGTATACAAGCGAAAGCAGCACTCAAGTAC 1

RESULT 12
US-09-801-274-517
; Sequence 517, Application US/09801274
; Patent No. US20020032119A1
; GENERAL INFORMATION:
; APPLICANT: Cargill, Michele
; APPLICANT: Ireland, James S.
; APPLICANT: Lander, Eric S.
; TITLE OF INVENTION: HUMAN SINGLE NUCLEOTIDE POLYMORPHISMS
; FILE REFERENCE: 2825.2009-001
; CURRENT APPLICATION NUMBER: US/09/801,274
; CURRENT FILING DATE: 2001-03-07
; PRIOR APPLICATION NUMBER: US 60/187,510
; PRIOR FILING DATE: 2000-03-07
; PRIOR APPLICATION NUMBER: US 60/206,129
; PRIOR FILING DATE: 2000-05-22
; NUMBER OF SEQ ID NOS: 1802
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 517
; LENGTH: 31
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-801-274-517

Query Match      1.1%  Score 19.2; DB 10; Length 31;
Best Local Similarity 80.8%; Pred. No. 3.9e+04;
Matches 21; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1190 CCACAGGCGGTCCCTCTTTCCGGGC 1215
   |||||  |||||  |||||  |||||  |||||
DB 5 CCACAGGCGTTCCCTCTTCGTGGGC 30

RESULT 13
US-09-263-959-121/c
; Sequence 121, Application US/09263959
; Patent No. US20020150891A1
; GENERAL INFORMATION:
; APPLICANT: Hood, Leroy E.
; APPLICANT: Rowen, Lee
; APPLICANT: Koop, Ben F.
; TITLE OF INVENTION: DIAGNOSTIC AND THERAPEUTIC COMPOSITIONS AND METHODS WHICH U
; NUMBER OF SEQUENCES: 1279
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Seed and Berry LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: US
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
```

```
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/263,959
; FILING DATE: 05-MAR-1999
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: McMaisters, David D.
; REGISTRATION NUMBER: 33,963
; REFERENCE/DOCKET NUMBER: 920010.426C2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 121:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 46 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-09-263-959-121
```

```
Query Match 1.1%; Score 19.2; DB 10; Length 46;
Best Local Similarity 75.0%; Pred. No. 4.7e+04;
Matches 24; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
```

```
QY 319 GCACCAGAGATTGTGCACGAGGACTTGAAGAT 350
||||| || ||||| ||| | |||||
Db 46 GCACCAGAGATTGTGCACGAGGAGGGGGAGAT 15
```

## RESULT 14

```
US-09-376-940-23/C
; Sequence 23, Application US/09376940A
; Publication No. US20020192813A1
; GENERAL INFORMATION:
; APPLICANT: Conner, Timothy W
; APPLICANT: Santino, Colleen G
; TITLE OF INVENTION: No. US20020192813A1el Plant Expression Vectors
; FILE REFERENCE: monocot elements
; CURRENT APPLICATION NUMBER: US/09/376,940A
; CURRENT FILING DATE: 1999-08-18
; EARLIER APPLICATION NUMBER: 60/097150
; EARLIER FILING DATE: 1998-08-19
; NUMBER OF SEQ ID NOS: 63
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 23
; LENGTH: 43
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:synthetic
; US-09-376-940-23
```

```
Query Match 1.1%; Score 19; DB 9; Length 43;
Best Local Similarity 65.1%; Pred. No. 5.1e+04;
Matches 28; Conservative 0; Mismatches 15; Indels 0; Gaps 0;
```

```
QY 79 GGGCCCGCGGCTCTGAGGTGCTGCGCGCGCGCGCGCGATCG 121
||||| | | | | | | | | | | | | | | | | | | | | |
Db 43 GTGTCCACCCACCTCGATCTCTGCTCGCGCGCGCGCGATCG 1
```

## RESULT 15

```
US-09-818-066-32
; Sequence 32, Application US/09818066
; Patent No. US20020032307A1
; GENERAL INFORMATION:
; APPLICANT: Shuping Tong et al.
; TITLE OF INVENTION: HEPADNAVIRUS RECEPTOR
; NUMBER OF SEQUENCES: 75
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
```

```
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/818,066
; FILING DATE: 27-Mar-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/683,262
; FILING DATE: 18-JUL-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Fraser, Janis K.
; REGISTRATION NUMBER: 31,819
; REFERENCE/DOCKET NUMBER: 00786/287002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-5070
; TELEFAX: (617) 542-8906
; TELEX: 200134
; INFORMATION FOR SEQ ID NO: 32:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 45 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-09-818-066-32
```

```
Query Match 1.1%; Score 19; DB 10; Length 45;
Best Local Similarity 71.4%; Pred. No. 5.2e+04;
Matches 25; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
```

```
QY 1343 CCCTTTTGAGCCAGCCGACCCGACTTGATAGCGAC 1377
||||| ||||| ||||| | | | | | | | | | | | | | | | | |
Db 2 CCCTTTTGAGCCAGCAATCCAGGATGAACACAGAC 36
```

```
Search completed: March 4, 2003, 00:54:25
Job time : 134 secs
```



DB 2 GCACGAATCCACCTGGCTGCCACTCCCTCTCTCTCTG 44

## RESULT 2

PCT-US95-13830-5  
: Sequence 5, Application PC/TUS9513830  
: GENERAL INFORMATION:  
: APPLICANT: Genentech, Inc.  
: TITLE OF INVENTION: Methods and Kits Using Macrophage Stimulating  
: TITLE OF INVENTION: Protein  
: NUMBER OF SEQUENCES: 10  
: CORRESPONDENCE ADDRESS:  
: ADDRESSEE: Genentech, Inc.  
: STREET: 460 Point San Bruno Blvd  
: CITY: South San Francisco  
: STATE: California  
: COUNTRY: USA  
: ZIP: 94080  
: COMPUTER READABLE FORM:  
: MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk  
: COMPUTER: IBM PC compatible  
: OPERATING SYSTEM: PC-DOS/MS-DOS  
: SOFTWARE: Winpatin (Genentech)  
: CURRENT APPLICATION DATA:  
: APPLICATION NUMBER: PCT/US95/13830  
: FILING DATE:  
: CLASSIFICATION:  
: ATTORNEY/AGENT INFORMATION:  
: NAME: Mairchang, Diane L.  
: REGISTRATION NUMBER: 35,600  
: REFERENCE/DOCKET NUMBER: P0912PCT  
: TELECOMMUNICATION INFORMATION:  
: TELEPHONE: 415/952-9881  
: TELEFAX: 415/225-5416  
: TELEX: 910/371-7168  
: INFORMATION FOR SEQ ID NO: 5:  
: SEQUENCE CHARACTERISTICS:  
: LENGTH: 47 base pairs  
: TYPE: Nucleic Acid  
: STRANDEDNESS: Single  
: TOPOLOGY: Linear  
PCT-US95-13830-5

Query Match 1.38; Score 22.2; DB 5; Length 47;  
Best Local Similarity 59.8%; Pred. No. 5.8e+03;  
Matches 30; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

OY 241 GCGCGCAGTGACCTGGAGAGGCCCCACACGCTGCT 283

DB 2 GCACGAATCCACCTGGCTGCCACTCCCTCTCTCTG 44

## RESULT 3

US-08-507-426C-9/c  
: Sequence 9, Application US/08507426C  
: Patent No. 6265634  
: GENERAL INFORMATION:  
: APPLICANT: Lenee, Philippe  
: TITLE OF INVENTION: POLYRIBOZYME CAPABLE OF CONFERRING ON PLANTS RESISTANCE  
: TITLE OF INVENTION: TO VIRUSES AND RESISTANT PLANTS PRODUCING THIS  
: TITLE OF INVENTION: POLYRIBOZYME  
: FILE REFERENCE: 43944-A-PCT-US  
: CURRENT APPLICATION NUMBER: US/08/507,426C  
: PRIOR FILING DATE: 1995-10-25  
: PRIOR APPLICATION NUMBER: 43944-A-PCT-US  
: PRIOR FILING DATE: 1995-10-25  
: NUMBER OF SEQ ID NOS: 14  
: SOFTWARE: Patent in Ver. 2.1  
: SEQ ID NO 9  
: LENGTH: 44  
: TYPE: DNA  
: ORGANISM: Artificial Sequence

: FEATURE:  
: OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
: OTHER INFORMATION: ribozymes and portions thereof  
US-08-507-426C-9

Query Match 1.28; Score 21.8; DB 4; Length 44;  
Best Local Similarity 70.7%; Pred. No. 7.2e+03;  
Matches 29; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

OY 1559 COTGGATCGCTGACTCAGGACGACGCTTCGCGCTG 1599

DB 42 COTGGTGGAGACTCAGGACGACGCTTCGCGCTG 2

## RESULT 4

US-08-171-389-156  
: Sequence 156, Application US/08171389  
: Patent No. 5578444  
: GENERAL INFORMATION:  
: APPLICANT: Edwards, Cynthia A.  
: APPLICANT: Cantor, Charles R.  
: APPLICANT: Andrews, Beth M.  
: APPLICANT: Turin, Lisa M.  
: APPLICANT: FTY, Kirk E.  
: TITLE OF INVENTION: Sequence-Directed DNA Binding  
: TITLE OF INVENTION: Molecules, Compositions and Methods  
: NUMBER OF SEQUENCES: 641  
: CORRESPONDENCE ADDRESS:  
: ADDRESSEE: Genelabs Technologies, Inc.  
: STREET: 505 Penobscot Drive  
: CITY: Redwood City  
: STATE: CA  
: COUNTRY: USA  
: ZIP: 94063  
: COMPUTER READABLE FORM:  
: MEDIUM TYPE: Floppy disk  
: COMPUTER: IBM PC compatible  
: OPERATING SYSTEM: PC-DOS/MS-DOS  
: SOFTWARE: Patent in Release #1.0, Version #1.25  
: CURRENT APPLICATION DATA:  
: APPLICATION NUMBER: US/08/171,389  
: FILING DATE:  
: CLASSIFICATION: 435  
: PRIOR APPLICATION DATA:  
: APPLICATION NUMBER: US 08/123,936  
: FILING DATE: 17-SEP-1993  
: PRIOR APPLICATION DATA:  
: APPLICATION NUMBER: US 07/995,783  
: FILING DATE: 23-DEC-1992  
: PRIOR APPLICATION DATA:  
: APPLICATION NUMBER: US 07/723,618  
: FILING DATE: 27-JUN-1991  
: PRIOR APPLICATION DATA:  
: APPLICATION NUMBER: US 08/081,070  
: FILING DATE: 21-JUN-1993  
: ATTORNEY/AGENT INFORMATION:  
: NAME: Fabian, Gary R.  
: REGISTRATION NUMBER: 33,875  
: REFERENCE/DOCKET NUMBER: 4600-0175/G19P3  
: TELECOMMUNICATION INFORMATION:  
: TELEPHONE: (415) 324-0880  
: TELEFAX: (415) 324-0960  
: INFORMATION FOR SEQ ID NO: 156:  
: SEQUENCE CHARACTERISTICS:  
: LENGTH: 46 base pairs  
: TYPE: nucleic acid  
: STRANDEDNESS: double  
: TOPOLOGY: linear  
: MOLECULE TYPE: DNA (genomic)  
: HYPOTHEetical: NO  
: ORIGINAL SOURCE:  
: INDIVIDUAL ISOLATE: Human H1 histone gene FNC16  
US-08-171-389-156





```

; APPLICANT: Cantor, Charles R.
; APPLICANT: Andrews, Beth M.
; APPLICANT: Turin, Lisa M.
; APPLICANT: Fry, Kirk E.
; TITLE OF INVENTION: Sequence-Directed DNA Binding
; TITLE OF INVENTION: Molecules, Compositions and Methods
; NUMBER OF SEQUENCES: 664
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genelabs Technologies, Inc.
; STREET: 505 Penobscot Drive
; CITY: Redwood City
; STATE: CA
; COUNTRY: USA
; ZIP: 94063
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/482,080A
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/171,389
; FILING DATE: 20-DEC-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/123,936
; FILING DATE: 17-SEP-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/996,783
; FILING DATE: 23-DEC-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/723,618
; FILING DATE: 27-JUN-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/081,070
; FILING DATE: 22-JUN-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Brady, John F.
; REGISTRATION NUMBER: 39,118
; REFERENCE/DOCKET NUMBER: 4600-0175.20/G19P3D1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 324-0880
; TELEFAX: (650) 324-0960
; INFORMATION FOR SEQ ID NO: 156:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 46 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; INDIVIDUAL ISOLATE: Human H1 histone gene PNC16
; US-08-482-080A-156

Query Match 1.28; Score 21.4; DB 3; Length 46;
Best Local Similarity 71.84; Pred. No. 9.3e+03;
Matches 28; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 1641 GCGGCTGGAGGATGCCACACCCCTCACAGGCGAGCC 1679
      |||||  |||||  |||||  |||||  |||||  |||||  ||
DB 2 GCGGTGGATTGGACGCTCCACCAATCACAGGCGAGCC 40

RESULT 8
US-09-354-947-156
; Sequence 156, Application US/09354947
; Patent No. 6384208
; GENERAL INFORMATION:
; APPLICANT: Edwards, Cynthia A.
; APPLICANT: Cantor, Charles R.
; APPLICANT: Andrews, Beth M.
; TITLE OF INVENTION: Sequence-Directed DNA Binding
; TITLE OF INVENTION: Molecules, Compositions and Methods
; NUMBER OF SEQUENCES: 641
```

```

; APPLICANT: Turin, Lisa M.
; APPLICANT: Fry, Kirk E.
; TITLE OF INVENTION: Sequence-Directed DNA Binding
; TITLE OF INVENTION: Molecules, Compositions and Methods
; NUMBER OF SEQUENCES: 664
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genelabs Technologies, Inc.
; STREET: 505 Penobscot Drive
; CITY: Redwood City
; STATE: CA
; COUNTRY: USA
; ZIP: 94063
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/354,947
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/482,080
; FILING DATE: 07-JUN-1995
; APPLICATION NUMBER: US 08/171,389
; FILING DATE: 20-DEC-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/123,936
; FILING DATE: 17-SEP-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/996,783
; FILING DATE: 23-DEC-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/723,618
; FILING DATE: 27-JUN-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/081,070
; FILING DATE: 22-JUN-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Brady, John F.
; REGISTRATION NUMBER: 39,118
; REFERENCE/DOCKET NUMBER: 4600-0175.20/G19P3D1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 324-0880
; TELEFAX: (650) 324-0960
; INFORMATION FOR SEQ ID NO: 156:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 46 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; INDIVIDUAL ISOLATE: Human H1 histone gene PNC16
; US-09-354-947-156

Query Match 1.28; Score 21.4; DB 4; Length 46;
Best Local Similarity 71.84; Pred. No. 9.3e+03;
Matches 28; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 1641 GCGGCTGGAGGATGCCACACCCCTCACAGGCGAGCC 1679
      |||||  |||||  |||||  |||||  |||||  |||||  ||
DB 2 GCGGTGGATTGGACGCTCCACCAATCACAGGCGAGCC 40

RESULT 9
PCT-US93-12388-156
; Sequence 156, Application PC/TUS9312388
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: Sequence-Directed DNA Binding
; TITLE OF INVENTION: Molecules, Compositions and Methods
; NUMBER OF SEQUENCES: 641
```





MEDIUM TYPE: FLOPPY DISK  
COMPUTER: IBM PC COMPATIBLE  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: WORDPERFECT 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/050,478  
FILING DATE: 26-OCT-1994  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/US93/07908  
FILING DATE: 29-MAR-1993  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/858,747  
FILING DATE: 27-MAR-1992  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: MORRY, MARY J.  
REGISTRATION NUMBER: 34,398  
REFERENCE/DOCKET NUMBER: 2026-40060S1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212)751-6849  
TELEFAX: (212)751-6849  
INFORMATION FOR SEQ ID NO: 47:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 48 BASE PAIRS  
TYPE: NUCLEIC ACID  
STRANDEDNESS: SINGLE  
TOPOLOGY: LINEAR  
US-08-050-478-47

Query Match 1.2%; Score 20.4; DB 2; Length 48;  
Best Local Similarity 71.1%; Pred. No. 1.7e+04;  
Matches 27; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

Oy 325 GAGATTGTGCACGAGGACTTGAAGATGGGCTCTGATCG 362  
||||| ||||| ||| ||||| ||||| ||||| ||||| |||||  
Db 7 GACACGCTGCCCTGAAGTTGAAGCCGGGGATGATCG 44

## RESULT 15

US-09-414-117-47  
Sequence 47, Application US/09414117  
Patent No. 6291664  
GENERAL INFORMATION:  
APPLICANT:  
APPLICANT:  
TITLE OF INVENTION: METHOD OF ELIMINATING  
INHIBITORY/INSTABILITY REGIONS OF mRNA  
NUMBER OF SEQUENCES: 130  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: MORGAN & FINNEGAN  
STREET: 345 PARK AVENUE  
CITY: NEW YORK  
STATE: NEW YORK  
COUNTRY: USA  
ZIP: 10154  
COMPUTER READABLE FORM:  
MEDIUM TYPE: FLOPPY DISK  
COMPUTER: IBM PC COMPATIBLE  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: WORDPERFECT 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/414,117  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/850,049  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/US93/02908

FILING DATE: 29-MAR-1993  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/858,747  
FILING DATE: 27-MAR-1992  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: MORRY, MARY J.  
REGISTRATION NUMBER: 34,398  
REFERENCE/DOCKET NUMBER: 2026-40060S1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212)751-6849  
TELEFAX: (212)751-6849  
INFORMATION FOR SEQ ID NO: 47:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 48 BASE PAIRS  
TYPE: NUCLEIC ACID  
STRANDEDNESS: SINGLE  
TOPOLOGY: LINEAR  
US-09-414-117-47

Query Match 1.2%; Score 20.4; DB 4; Length 48;  
Best Local Similarity 71.1%; Pred. No. 1.7e+04;  
Matches 27; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

Oy 325 GAGATTGTGCACGAGGACTTGAAGATGGGCTCTGATCG 362  
||||| ||||| ||| ||||| ||||| ||||| ||||| |||||  
Db 7 GACACGCTGCCCTGAAGTTGAAGCCGGGGATGATCG 44

Search completed: March 4, 2003, 00:52:00  
Job time : 80 secs

GenCore version 5.1.4.p5.4378  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run On: March 3, 2003, 19:29:35 : Search time 427 Seconds  
(without alignments)  
9203.131 Million cell updates/sec

Title: US-10-017-621-3  
Perfect score: 1745  
Sequence: 1 tggagacagcgttaagatg.....gtccactcccaactgtcc 1745

Scoring table: IDENTITY\_NUC  
Gapop 10.0, Gapext 1.0

Searched: 2185239 seqs, 112599159 residues  
Total number of hits satisfying chosen parameters: 2156140

Minimum DB seq length: 0  
Maximum DB seq length: 50

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :	N_Geneseq_101002:*
1:	/SID52/gcgdata/geneseq/geneseq-emb1/NA1980.DAT:*
2:	/SID52/gcgdata/geneseq/geneseq-emb1/NA1981.DAT:*
3:	/SID52/gcgdata/geneseq/geneseq-emb1/NA1982.DAT:*
4:	/SID52/gcgdata/geneseq/geneseq-emb1/NA1983.DAT:*
5:	/SID52/gcgdata/geneseq/geneseq-emb1/NA1984.DAT:*
6:	/SID52/gcgdata/geneseq/geneseq-emb1/NA1985.DAT:*
7:	/SID52/gcgdata/geneseq/geneseq-emb1/NA1986.DAT:*
8:	/SID52/gcgdata/geneseq/geneseq-emb1/NA1987.DAT:*
9:	/SID52/gcgdata/geneseq/geneseq-emb1/NA1988.DAT:*
10:	/SID52/gcgdata/geneseq/geneseq-emb1/NA1989.DAT:*
11:	/SID52/gcgdata/geneseq/geneseq-emb1/NA1990.DAT:*
12:	/SID52/gcgdata/geneseq/geneseq-emb1/NA1991.DAT:*
13:	/SID52/gcgdata/geneseq/geneseq-emb1/NA1992.DAT:*
14:	/SID52/gcgdata/geneseq/geneseq-emb1/NA1993.DAT:*
15:	/SID52/gcgdata/geneseq/geneseq-emb1/NA1994.DAT:*
16:	/SID52/gcgdata/geneseq/geneseq-emb1/NA1995.DAT:*
17:	/SID52/gcgdata/geneseq/geneseq-emb1/NA1996.DAT:*
18:	/SID52/gcgdata/geneseq/geneseq-emb1/NA1997.DAT:*
19:	/SID52/gcgdata/geneseq/geneseq-emb1/NA1998.DAT:*
20:	/SID52/gcgdata/geneseq/geneseq-emb1/NA1999.DAT:*
21:	/SID52/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:*
22:	/SID52/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:*
23:	/SID52/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:*
24:	/SID52/gcgdata/geneseq/geneseq-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	24.8	1.4	50	22	AA134335
C 2	22.4	1.3	33	24	ABA04099
C 3	22.4	1.3	33	24	ABA04100
C 4	22	1.3	31	22	AA130264
C 5	22	1.3	50	22	AA175839
C 6	21.4	1.2	31	22	AA129606
C 7	21.4	1.2	42	22	AAH22523
C 8	21.4	1.2	42	24	AAH29563
C 9	21.4	1.2	46	15	AAQ69406

10	21.4	1.2	46	18	AAT63868	Human H1 histone 9
11	21.4	1.2	46	20	AA17156	Test sequence from
12	21.4	1.2	46	24	ABK82647	DNA binding molecu
13	21	1.2	46	24	AAH62195	CTA18G-1 polymorp
C 14	21	1.2	46	22	AAH70659	Human cervical can
C 15	21	1.2	46	22	AA134286	Human SNP oligonuc
C 16	20.8	1.2	46	20	AA132932	DE19736531 primer
C 17	20.6	1.2	45	16	AA170598	RT-PCR primer/prob
C 18	20.6	1.2	45	16	AA170670	Primer 143 for hum
C 19	20.6	1.2	45	16	AA172077	Human SNP oligonuc
C 20	20.4	1.2	32	24	AA172077	Human SNP oligonuc
C 21	20.4	1.2	48	14	AAQ50230	Xcds1 degenerate p
C 22	20.2	1.2	36	19	AAV46356	HIV pol INS mutag
C 23	20.2	1.2	40	16	AAQ76190	PCR primer for ser
C 24	20.2	1.2	40	18	AA191033	Primer for amplif
C 25	20.2	1.2	40	24	ABL54053	Human 4-1BB 3' PCR
C 26	20.2	1.2	40	24	ABK48669	Human cytokine rec
C 27	20.2	1.2	41	24	ABK48669	Human proton-aden
C 28	20.2	1.2	41	24	ABK48669	Human proton-aden
C 29	20.2	1.2	45	22	AA17287	MUSIGHAEI Mouse Ig
C 30	20.2	1.2	49	20	AAZ31379	Reverse PCR primer
C 31	20.2	1.2	50	19	AAV59127	Reverse PCR primer
C 32	20	1.1	50	22	AAZ3538	CD4+ human T-lymph
C 33	20	1.1	41	19	AAV37843	Human large protei
C 34	20	1.1	46	21	AAZ91309	IL-6R and IL-6 fus
C 35	20	1.1	47	21	AAZ67885	Human map-related
C 36	20	1.1	50	22	AA129783	Human SNP oligonuc
C 37	20	1.1	50	22	AA134645	Human SNP oligonuc
C 38	19.8	1.1	46	21	AA134645	Human SNP oligonuc
C 39	19.8	1.1	48	18	AA170636	Hammerhead ribozym
C 40	19.8	1.1	48	20	AAZ53859	Human A2B adenosin
C 41	19.8	1.1	48	20	AAZ53859	Human adenosine A2
C 42	19.8	1.1	48	21	AA19424	Human adenosine A2
C 43	19.8	1.1	48	21	AAZ33302	Low adenosine anti
C 44	19.8	1.1	48	21	AAZ33302	Human adenosine A1
C 45	19.6	1.1	31	22	AAH74232	Nucleotide sequenc
						Human single nucle

# ALIGNMENTS

RESULT 1	
AA134335/c	
ID	AA134335 standard; DNA; 50 BP.
XX	AA134335;
XX	24-JAN-2002 (first entry)
XX	Human SNP oligonucleotide #7543.
XX	Immunosuppressive; immunostimulatory; antiinflammatory; cytostatic;
XX	neuroprotective; antimicrobial; gene therapy; vaccine; amylose; cancer;
XX	anyoid protein; angiotensin; apoptosis related protein; cadherin;
XX	cyclin; polymerase; oncogene; histone; kinase; colony stimulating factor;
XX	complement related protein; cytochrome; kinesin; cytokine; interferon;
XX	interleukin; G-protein coupled receptor; thioesterase; inflammation;
XX	multifactorial disease; autoimmune disease; infection;
XX	nervous system disease; ss.
OS	Homo sapiens.
XX	WO200147944-A2.
PN	05-JUL-2001.
PD	28-DEC-2000; 2000WO-US35498.
XX	28-DEC-1999; 99US-0173419.
XX	27-DEC-2000; 2000US-0173419.
XX	(CURA-) CURAGEN CORP.

The present invention describes medicinal compositions (I) inhibiting beta-amyloid production comprising an active component a substance that inhibits the activity of cyclin-dependent kinase (CDK). Also described are: (1) a method for screening compounds for their ability to inhibit the production of beta-amyloid by contacting with beta-amyloid producing cells; and (2) screening kits. (1) have neurotropic and neuroprotective activities. (1) suppress the phosphorylation of amyloid precursor protein (APP) which is an essential step in the production of beta-amyloid. (1) can be used in the treatment and prevention of neurodegenerative diseases such as dementia and Alzheimer's disease. The present sequence represents a PCR primer which is used in the exemplification of the present invention.

Sequence 33 BP; 6 A; 6 C; 11 G; 10 T; 0 other;

Query Match 1.38; Score 22.4; DB 24; Length 33;  
Best Local Similarity 81.28; Pred. NO. 2.8e+04;  
Matches 26; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 1018 GAGCTCAACGTCGGTGTGCTTGCCCTGGCGCCG 1049  
||||| || ||||| || ||||| || ||||| ||  
DB 2 GAGCTGAATTTGGCTAAATTTGGCCTGCCTCG 33

RESULT 3  
ABA04100/C  
ID ABA04100 standard; DNA; 33 BP.  
XX  
AC ABA04100;  
XX  
DT 21-FEB-2002 (first entry)  
DE Human Cdk5 related PCR primer SEQ ID NO:19.  
XX  
KW Human; beta-amyloid; cyclin-dependent kinase inhibitor; nerve cell;  
KW amyloid precursor protein; APP; Cdk5; PCR primer; ss.  
OS Homo sapiens.  
XX  
PN W0200182967-A1.  
XX  
PD 08-NOV-2001.  
XX  
PF 25-APR-2001; 2001WO-JP03555.  
XX  
PR 28-APR-2000; 2000JP-0131037.  
XX  
PA {YAMA } YAMANOUCHI PHARM CO LTD.  
PA {SUZU/} SUZUKI T.  
XX  
PI Suzuki T., Watanabe T., Kawabata S., Hachiya S.;  
DR WPI; 2002-026209/03.  
XX  
PT Medicinal compositions for the treatment of dementia and Alzheimer's disease, comprise compounds that suppress beta amyloid production - Example 6; Page 23; 62pp; Japanese.  
XX  
The present invention describes medicinal compositions (I) inhibiting beta-amyloid production comprising an active component a substance that inhibits the activity of cyclin-dependent kinase (CDK). Also described are: (1) a method for screening compounds for their ability to inhibit the production of beta-amyloid by contacting with beta-amyloid producing cells; and (2) screening kits. (1) have neurotropic and neuroprotective activities. (1) suppress the phosphorylation of amyloid precursor protein (APP) which is an essential step in the production of beta-amyloid. (1) can be used in the treatment and prevention of neurodegenerative diseases such as dementia and Alzheimer's disease. The present sequence represents a PCR primer which is used in the exemplification of the present invention.

Sequence 33 BP; 10 A; 11 C; 6 G; 6 T; 0 other;

Shinkets RA, Leach M;  
WPI; 2001-465210/50.  
Polymorphic nucleic acids encoding e.g. amylases, cyclins, polymerases, oncogenes and histones, useful for diagnosing and treating, e.g. cancer, autoimmune diseases and infections .  
Claim 1; Page 3563; 4143pp; English.  
The present invention relates to oligonucleotides encoding polymorphic variants of proteins related to amylases, amyloid proteins, angiotensin, apoposis related proteins, cadherin, cyclin, polymerase, oncogenes, histones, kinases, colony stimulating factors, complement related proteins, cytochromes, kinesins, cytokines, interferons, interleukins, G-protein coupled receptors and cholesterales. The present sequence is one such oligonucleotide. The oligonucleotides and the peptides encoded by them may be used in the prevention, diagnosis and treatment of diseases associated with inappropriate expression of the proteins listed above. Disorders that may be prevented, diagnosed and/or treated include multifactorial diseases with a genetic component, such as autoimmune diseases (e.g. rheumatoid arthritis, multiple sclerosis, diabetes, systemic lupus erythematosus and Grave's disease), inflammation, cancer (e.g. cancers of the bladder, brain, breast, colon and kidney, leukaemia), diseases of the nervous system and an infection of pathogenic organisms.  
Sequence 50 BP; 7 A; 13 C; 14 G; 16 T; 0 other;  
Query Match 1.48; Score 24.8; DB 22;  
Best Local Similarity 72.78; Pred. No. 7.8e-03;  
Matches 32; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

OY 1619 CAGACCGAGGCCCGACGAGCGGCTGGGAGGTCGCCACACC 1652  
||||| ||||| ||||| ||||| ||||| || |||||  
DB 45 CAGACCGAGGCCCGACGAGCTCATGCTGGAGAATTATTCACACC 2

RESULT 2  
ABA04099  
ID ABA04099 standard; DNA; 33 BP.  
XX  
AC ABA04099;  
XX  
DT 21-FEB-2002 (first entry)  
DE Human Cdk5 related PCR primer SEQ ID NO:18.  
XX  
KW Human; beta-amyloid; cyclin-dependent kinase inhibitor; nerve cell;  
KW amyloid precursor protein; APP; Cdk5; PCR primer; ss.  
OS Homo sapiens.  
XX  
PN W0200182967-A1.  
XX  
PD 08-NOV-2001.  
XX  
PF 25-APR-2001; 2001WO-JP03555.  
XX  
PR 28-APR-2000; 2000JP-0131037.  
XX  
PA {YAMA } YAMANOUCHI PHARM CO LTD.  
PA {SUZU/} SUZUKI T.  
XX  
PI Suzuki T., Watanabe T., Kawabata S., Hachiya S.;  
DR WPI; 2002-026209/03.  
XX  
PT Medicinal compositions for the treatment of dementia and Alzheimer's disease, comprise compounds that suppress beta amyloid production - Example 6; Page 23; 62pp; Japanese.







XX DR WPI; 2002-171810/22.  
 XX PS  
 XX CC Producing a protein of interest, e.g., a pharmaceutically active  
 PT protein, comprises expressing a polynucleotide fusion construct in a  
 PT plasmid and producing a fusion protein comprising the protein of  
 PT interest.  
 XX CC  
 XX PS Example 1; Page 90; 92pp; English.  
 XX CC  
 CC The patent discloses a method of producing a protein of interest which  
 CC involves expressing a polynucleotide fusion construct in a plasmid to  
 CC produce a fusion protein comprising the protein of interest where the  
 CC construct comprises a polynucleotide coding sequence of the protein of  
 CC interest operably linked to a polynucleotide coding sequence of a fusion  
 CC protein partner. The methods of the invention are useful for producing a  
 CC protein of interest which comprises a human protein or its biologically  
 CC active variant or fragment, a pharmaceutically active protein, an IFN-  
 CC (interferon), its biologically active variant or fragment, a human IFN-  
 CC gamma or its biologically active variant or fragment. They are useful  
 CC for the production of transgenic plants. Methods of the invention are  
 CC also useful for the generation of transplasmatic plant cells, plants  
 CC and seeds. The protein of interest obtained by the methods of the  
 CC invention is useful for the manufacture of a medicament for treating  
 CC a disease condition. The present DNA sequence is a PCR primer which  
 CC is used for amplifying ifnc coding region. This primer is used in  
 CC the exemplification of the invention to generate pcUSIFNG expression  
 CC vector for ifnc in chloroplasts.  
 XX CC  
 XX SQ Sequence 42 BP; 15 A; 10 C; 11 G; 6 T; 0 other;  
 Query Match 1.24; Score 21.4; DB 24; Length 42;  
 Best Local Similarity 80.6%; Pred. No. 5.6e+04;  
 Matches 25; Conservative 0; Mismatches 6; Indels 0; Gaps 0;  
 OY 270 ACGTGTGCTCTCGGGCAACTGCTTCGCA 300  
 DB 35 ACGTACGGTCTCTGGCAGCTTCGATCGCA 5  
 RESULT 9  
 AAQ69406  
 ID AAQ69406 standard; DNA: 46 BP.  
 XX AC AAQ69406;  
 XX DT 27-FEB-1995 (first entry)  
 XX DE Human H1 histone gene FNC16, target region.  
 XX KW DNA protein-binding assay; test sequence; screening sequence;  
 KW promoter; target; TATA box; Herpes Simplex Virus; HSV;  
 KW origin of replication; UL9; transcription factor; TFIIID; ds.  
 XX OS Synthetic.  
 XX PN WO9414980-A.  
 XX PD 07-JUL-1994.  
 XX PF 20-DEC-1993; 93WO-US12388.  
 XX PR 23-DEC-1992; 92US-0996783.  
 PR 17-SEP-1993; 93US-0123936.  
 XX PA (GENE-) GENELABS TECHNOLOGIES INC.  
 XX PI Andrews BM, Cantor CR, Edwards CA, Fry KE, Turin LM;  
 XX DR WPI; 1994-234711/28.  
 XX CC Sequence-directed DNA-binding molecules - useful in  
 PT pharmaceuticals and as molecular reagents

XX PS Claim 28; Page 290; 587pp; English.  
 XX CC A DNA protein-binding assay is provided, useful for screening  
 CC libraries of synthetic or biological cpds. for their ability  
 CC to bind DNA test sequences. The assay is versatile in that any  
 CC number of test sequences can be tested by placing the test sequence  
 CC adjacent to a defined protein-binding screening sequence. Binding  
 CC of mols. to these test sequences changes the binding characteristics  
 CC of the protein mol. to its cognate binding sequence. When such a mol.  
 CC binds the test sequence, the equilibrium of the DNA:protein complex  
 CC is disturbed, generating changes in the concentration of free DNA probe.  
 CC One application of this method is to eucaryotic general transcription  
 CC factors (e.g. TFIID) where the target region is typically selected  
 CC from DNA sequences adjacent to the binding site for the eucaryotic  
 CC transcription factor. Numerous exemplary test sequences are given:  
 CC the sequences in AAQ69251-731 and AAQ69850 correspond to promoter  
 CC targets (typically, TATA box-contg. sites) for human genes and the  
 CC sequences in AAQ69732-849 correspond to promoter targets for viral genes.  
 CC The test sequences may also be randomly generated. DNA:protein  
 CC interaction may be used for screening purposes, e.g. the Herpes Simplex  
 CC Virus (HSV) origin of replication and UL9 (see AAQ69851-52, AAQ69865 and  
 CC AAQ69891).  
 XX CC  
 XX SQ Sequence 46 BP; 9 A; 14 C; 16 G; 7 T; 0 other;  
 Query Match 1.24; Score 21.4; DB 15; Length 46;  
 Best Local Similarity 71.8%; Pred. No. 5.8e+04;  
 Matches 28; Conservative 0; Mismatches 11; Indels 0; Gaps 0;  
 OY 1641 GCGGTGAGGGGATGCCACACCCCTCACAGGCGAGCC 1679  
 DB 2 GCGGTGATGGACGCTCCACCAATCACAGGCGAGCC 40  
 RESULT 10  
 AAQ63868  
 ID AAQ63868 standard; DNA: 46 BP.  
 XX AC AAQ63868;  
 XX DT 14-MAR-1997 (first entry)  
 XX DE Human H1 histone gene FNC16 gene TFIID binding site.  
 XX KW Duplex DNA; target region; binding characteristic; DNA binding protein;  
 KW TFIID; transcription factor; binding site; inhibition; enhance;  
 KW cancer; inherited genetic disorder; alpha-D-galactosidase A; ds.  
 XX OS Homo sapiens.  
 XX PN US5578444-A.  
 XX PD 26-NOV-1996.  
 XX PF 27-JUN-1991; 91US-0723618.  
 XX PR 20-DEC-1993; 93US-0171389.  
 PR 27-JUN-1991; 91US-0723618.  
 PR 23-DEC-1992; 92US-0996783.  
 PR 17-SEP-1993; 93US-0123936.  
 XX PA (GENE-) GENELABS TECHNOLOGIES INC.  
 XX PI Andrews BM, Cantor CR, Edwards CA, Fry KE, Turin LM;  
 XX DR WPI; 1997-020402/02.  
 XX CC Altering binding characteristics of DNA binding proteins to duplex  
 PT DNA - by attaching specific small cpd. to target region close to the  
 PT protein's binding site, useful in treatment of viral disease, cancer  
 PT etc

PS Claim 6: Column 177-178; 264pp; English.

XX The sequences given in AAT63713-4312 represent duplex DNA's which act  
CC as target regions in the method of the invention. The method for  
CC altering the binding characteristics of a DNA-binding protein to duplex  
CC DNA comprises contacting the duplex DNA with a small molecule which  
CC binds sequence-specifically to a target region, where, when the small  
CC molecule is bound to the target region, it is adjacent to, but not  
CC overlapping by more than 4 bp, a binding site for a DNA-binding protein.  
CC The small molecule is added at a concentration effective to alter the  
CC binding of the DNA binding protein, pref. TFIID, to its binding site on  
CC the duplex DNA. The binding of the small molecule may inhibit or  
CC enhance the binding of the DNA-binding protein to its binding site. The  
CC compounds isolated using this method are potentially useful as  
CC therapeutic agents for treatment of any disease which involves a  
CC specific DNA sequence, e.g. cancer, or inherited genetic disorders etc.  
CC The method is suitable for screening large biological or chemical  
CC libraries and allows determination of sequence-specific and relative  
CC affinities of known DNA-binding agents for different DNA sequences.  
CC The design of these duplex DNA's allows a single DNA:protein interaction  
CC to be used for screening sequence-specific, or preferential, DNA binding  
CC proteins that recognise almost any possible sequence (see also AAT49539-  
CC 74).

XX Sequence 46 BP; 9 A; 14 C; 16 G; 7 T; 0 other;

Query Match 1.28; Score 21.4; DB 18; Length 46;

Best Local Similarity 71.88; Pred. No. 5.8e+04;

Matches 28; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 1641 GCGGCTGAGGATGCCACACCCCTCACAGGCGAGCC 1679

DB 2 GCGGTGGATTGGACGCTCCACCAATCACAGGCGAGCC 40

RESULT 11

AAAX17156

ID AAX17156 standard; DNA: 46 BP.

XX AAX17156;

XX 06-MAY-1999 (first entry)

XX Test sequence from human H1 histone gene FNCL6.

XX Test sequence; DNA-binding molecule; screening sequence; human;

XX nucleic acid amplification; target; viral; ds.

XX Homo sapiens.

XX U55869241-A.

XX 09-FEB-1999.

XX 07-JUN-1995; 95US-0475228.

XX 20-DEC-1993; 93US-0171389.

XX 27-JUN-1991; 91US-0723618.

XX 23-DEC-1992; 92US-0996783.

XX 17-SEP-1993; 93US-0123936.

XX 07-JUN-1995; 95US-0475228.

XX (GENE-) GENELABS TECHNOLOGIES INC.

XX Andrews BM, Cantor CR, Edwards CA, Fry KE, Turin LM;

XX WPI; 1999-152755/13.

XX Determination of DNA sequence preference of a DNA-binding molecule -  
PT based on inhibition of binding of protein to oligonucleotide  
PT sequence attached to test sequence

XX Claim 3; Columns 179-180; 270pp; English.

XX

CC Sequences AAX17001 to AAX17600 represent specifically claimed target  
CC test sequences that are used in the method of the invention of  
CC determining the DNA sequence preference of a DNA-binding molecule. The  
CC method comprises: (i) adding a test molecule and a DNA-binding protein to  
CC a mixture of duplex DNA test oligonucleotides, each of the test  
CC oligonucleotides having a test sequence adjacent to a screening sequence,  
CC where the screening sequence binds to the DNA-binding protein with a  
CC binding affinity that is independent of the DNA sequence of the test  
CC sequence, and where the mixture of duplex DNA test oligonucleotides  
CC includes several test sequences; (ii) incubating the test molecule, the  
CC mixture of duplex DNA test oligonucleotides and the DNA-binding protein  
CC for a time sufficient to permit binding of the test molecule to test  
CC sequences in the duplex DNA; (iii) separating unbound test  
CC oligonucleotides from test oligonucleotides bound to binding proteins;  
CC (iv) amplifying the unbound test oligonucleotides; (v) repeating steps  
CC (ii) to (iv); (vi) isolating the amplified test oligonucleotides; and  
CC (vii) sequencing the isolated test oligonucleotides. Test sequences  
CC AAX17001-X17481 and AAX17600 correspond to promoter targets for human  
CC genes and test sequences AAX17482-X17599 correspond to promoter targets  
CC for viral genes.

XX Sequence 46 BP; 9 A; 14 C; 16 G; 7 T; 0 other;

Query Match 1.28; Score 21.4; DB 20; Length 46;

Best Local Similarity 71.88; Pred. No. 5.8e+04;

Matches 28; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 1641 GCGGCTGAGGATGCCACACCCCTCACAGGCGAGCC 1679

DB 2 GCGGTGGATTGGACGCTCCACCAATCACAGGCGAGCC 40

RESULT 12

ABK82647

ID ABK82647 standard; DNA: 46 BP.

XX AC ABK82647;

XX 27-AUG-2002 (first entry)

XX DNA binding molecule screening method test sequence #156.

XX DNA binding molecule screening; inhibition of transcription;  
KW infection; human immunodeficiency virus; HIV; parasite; cancer;  
KW cardiovascular; respiratory; gastrointestinal; endocrine; metabolic;  
KW rheumatic; immunological; haematological; neurological;

KW psychiatric; dermatological; ophthalmological; musculo-skeletal;  
KW urogenital disorder; ss.

XX Synthetic.

XX US6384208-B1.

XX 07-MAY-2002.

XX 15-JUL-1999; 99US-0354947.

XX 20-DEC-1993; 93US-0171389.

XX 07-JUN-1995; 95US-0482080.

XX 27-JUN-1991; 91US-0723618.

XX 23-DEC-1992; 92US-0996783.

XX 17-SEP-1993; 93US-0123936.

XX (GENE-) GENELABS TECHNOLOGIES INC.

XX Edwards CA, Cantor CR, Andrews BM, Turin LM, Fry KE;

XX WPI; 2002-442819/47.

XX Decreasing transcriptional activity of genes for treating infections or  
PT cancer, by administration of an agent that binds to two non-overlapping  
PT regions of the gene

CC synovial, connective tissue, rheumatoid arthritis, immunological, haematological,  
CC neurological, psychiatric, dermatological, ophthalmological,  
CC musculo-skeletal, genetic or urogenital disorders. The method provides  
CC sequence-specific inhibition of transcription of pathological genes  
CC without affecting transcription of cellular genes regulated by the same  
CC transcription factor, and can be applied to regulation of any gene.  
CC AKB82492-ADK9315 represent DNA binding molecule test sequences used in  
CC the method of the invention.

SQ Sequence 21 BP; 9 A; 4 C; 6 G; 2 T; 0 other;  
 Query Match 1.2A; Score 21; DB 22; Length 21;

Best Local Similarity 100.00; Pred. No. 5.4e+04;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Caps 0;

QY 702 CAAGGAGATCAGACTGGAACA 722  
|||||  
Db 1 CAAGGAGATCAGACTGGAACA 21

**RESULT 14**

AAH70659/c  
ID AAH70659 standard: cDNA: 46 BP.

XX  
AC  
AAH70659;

XX  
DT 19-SEP-2001 (first entry)

XX  
DE  
Human cervical cancer marker nucleic acid 1933.

XX  
KW  
v  
Cervical cancer; cytostatic; pre-malignant condition; gene therapy; ss.

YY  
OS  
Homo sapiens.

XX  
PN WO200142467-A2.

XX  
PD 14-JUN-2001

XX  
PF 08-DEC-2000:

XX  
PR

PR	21-DEC-1999;	99US-0171350.
PR	14-MAR-2000.	2000US 0180212

12-MAY-2000; 2000US-0203791.

PR 09-JUN-2000; 2000US-0210600.  
PR 21-JUL-2000; 2000US-0220114.

XX  
PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

XX  
PI Schlegel R, Deeds J, Berger A, Zhao X.

XX  
DR WPI: 2001-375006/39

xx  
pt  
pt  
xx  
New isolated nucleic acid for diagnosing and treating cervical cancer and for assessing and detecting compounds for treating the cancer.

PS Claim 1; Page 415; 1051pp; English.

The invention relates to novel genes (AAH58737-AAH73383) associated with cervical cancer with cytostatic activity. The nucleic acids and encoded polypeptides are useful to assess if a patient is afflicted with cervical cancer or has a pre-malignant condition; to monitor the progression of cervical cancer or a premalignant condition in a patient; and to select and/or assess the efficacy of a compound or therapy for inhibiting cervical cancer in a patient. The nucleic acids may also be useful for gene therapy.

Sequence 46 BP; 7 A; 13 C; 18 G; 6 T; 2 other;  
SSQ



GenCore version 5.1.4.p5.4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sv model

Run on: March 3, 2003, 21:30:10 ; Search time 4788 Seconds  
(without alignments)  
10606.599 Million cell updates/sec

Title: US-10-017-621-3  
Perfect score: 1745  
Sequence: 1 tggagcagcgttaagagatg.....gttcacactgcccactgtctc 1745

Scoring table: IDENTITY\_NUC  
Gap 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues  
Total number of hits satisfying chosen parameters: 841850

Minimum DB seq length: 0  
Maximum DB seq length: 50

Post-processing: Minimum Match 0  
Maximum Match 100  
Listing first 45 summaries

Database : GenBank

- 1: gb.ba.\*
- 2: gb.htg.\*
- 3: gb.in.\*
- 4: gb.om.\*
- 5: gb.ov.\*
- 6: gb.pat.\*
- 7: gb.ph.\*
- 8: gb.pl.\*
- 9: gb.pr.\*
- 10: gb.ro.\*
- 11: gb.sts.\*
- 12: gb.sy.\*
- 13: gb.un.\*
- 14: gb.vi.\*
- 15: em.ba.\*
- 16: em.fun.\*
- 17: em.hum.\*
- 18: em.in.\*
- 19: em.mu.\*
- 20: em.om.\*
- 21: em.or.\*
- 22: em.ov.\*
- 23: em.pat.\*
- 24: em.ph.\*
- 25: em.pl.\*
- 26: em.ro.\*
- 27: em.sts.\*
- 28: em.un.\*
- 29: em.vi.\*
- 30: em.htg.hum.\*
- 31: em.htg.in.\*
- 32: em.htg.other.\*
- 33: em.htg.mus.\*
- 34: em.htg.pln.\*
- 35: em.htg.rod.\*
- 36: em.htg.mam.\*
- 37: em.htg.vrt.\*
- 38: em.sy.\*
- 39: em.htgo.hum.\*
- 40: em.htgo.mus.\*
- 41: em.htgo.other.\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result	No.	Score	Query	Match	Length	DB	ID	Description
C	1	23.8	1.4	49	9	HSHCAK	X76171	H.sapiens m
	2	22.8	1.3	50	6	A93721	A93721	Sequence 8
	3	22.4	1.3	50	6	A93722	A93722	Sequence 9
	4	22.2	1.3	47	6	I84671	I84671	Sequence 5
	5	22	1.3	50	6	AX159452	AX159452	Sequence 5
	6	21.6	1.2	31	6	AX248673	AX248673	Sequence
C	7	21.4	1.2	42	6	AX182243	AX182243	Sequence
	8	21.4	1.2	42	6	AX382049	AX382049	Sequence
C	9	21.4	1.2	46	6	AR032544	AR032544	Sequence
	10	21.4	1.2	46	6	AR209208	AR209208	Sequence
	11	21.4	1.2	46	6	I29284	I29284	Sequence 15
	12	21.4	1.2	46	6	I30958	I30958	Sequence 15
	13	21	1.2	31	6	AX248015	AX248015	Sequence
C	14	21	1.2	46	6	AX186238	AX186238	Sequence
C	15	20.8	1.2	46	6	A96791	A96791	Sequence 24
	16	20.6	1.2	21	6	AX153998	AX153998	Sequence
C	17	20.6	1.2	45	6	AR022074	AR022074	Sequence
	18	20.6	1.2	45	6	I55009	I55009	Sequence 13
C	19	20.6	1.2	45	6	I32864	I32864	Sequence 18
	20	20.4	1.2	48	6	AR079723	AR079723	Sequence
	21	20.4	1.2	48	6	AR081253	AR081253	Sequence
	22	20.4	1.2	48	6	AR170613	AR170613	Sequence
	23	20.2	1.2	40	6	AR200128	AR200128	Sequence
	24	20.2	1.2	40	6	I88030	I88030	Sequence 13
C	25	20.2	1.2	43	6	AX225269	AX225269	Sequence
	26	20.2	1.2	49	6	AR083818	AR083818	Sequence
	27	20.2	1.2	49	9	S82032	S82032	Wtl-Wilms'
C	28	20.2	1.2	50	6	AX233404	AX233404	Sequence
	29	20	1.1	36	6	A07324	A07324	Synthetic D
	30	20	1.1	36	6	I12501	I12501	Sequence 18
	31	20	1.1	41	6	BD007098	BD007098	Targeted
C	32	20	1.1	44	6	A07325	A07325	Synthetic D
	33	20	1.1	44	6	I12502	I12502	Sequence 19
	34	20	1.1	46	6	E52011	E52011	IL-6 recept
	35	19.8	1.1	46	6	AX036348	AX036348	Sequence
	36	19.8	1.1	46	6	AX036350	AX036350	Sequence
C	37	19.8	1.1	48	6	BD012118	BD012118	Vitamin D
C	38	19.8	1.1	48	23	BD004595	BD004595	Vitamin D
	39	19.6	1.1	39	6	AX452342	AX452342	Sequence
C	40	19.6	1.1	42	6	AR153233	AR153233	Sequence
	41	19.6	1.1	45	6	I17261	I17261	Sequence 27
C	42	19.6	1.1	45	6	AR032970	AR032970	Sequence 19
	43	19.6	1.1	50	6	AR209634	AR209634	Sequence
	44	19.6	1.1	50	6	AX199648	AX199648	Sequence
	45	19.6	1.1	50	6	AX199648	AX199648	Sequence

ALIGNMENTS

RESULT	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL
1	HSHCAK	H.sapiens mRNA for Cdk activating kinase.	X76171	GI:429096	activating kinase; protein kinase.	Homo sapiens.	Homo sapiens	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.	1 (bases 1 to 49)	Hall, F.L.	Direct Submission
										Submitted (08-NOV-1993)	F.L. Hall, Childrens Hospital Los Angeles,



REFERENCE	1 (PAGES 1 TO 42)
AUTHORS	Reddy, S. I., Sadhu, L. I., Shukla, V. C. and Ferralolo, G. I.
TITLE	Plastid transformation





BASE COUNT 7 a 13 c 18 g 6 t 2 others  
ORIGIN /organism="Homo sapiens"  
/db\_xref="taxon:9606"

Query Match 1.28; Score 21; DB 6; Length 46;  
Best Local Similarity 71.1%; Pred. No. 3.5e+06;  
Matches 27; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

Oy 550 AAGCCCTCAGCGCGCGCTCGTGTGAGCTATC 587  
||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 41 AAGCGTCTGTGACGCGCGCGCTCGTGTGATC 4

## RESULT 15

A98791/c A98791 46 bp DNA linear PAT 26-JAN-2000  
LOCUS

DEFINITION Sequence 24 from Patent WO9910358.

ACCESSION A98791

VERSION A98791.1 GI:6781812

KEYWORDS

SOURCE unidentified.

ORGANISM unclassified.

REFERENCE 1 (bases 1 to 46)

AUTHORS Hegemann,P.

TITLE METHOD FOR PRODUCING NUCLEIC ACID POLYMERS

JOURNAL Patent: WO 9910358-A 24 04-MAR-1999;

HEGEMANN PETER (DE)

FEATURES

source

1..46

/organism="unidentified"

/db\_xref="taxon:32644"

BASE COUNT 9 a 13 c 15 g 8 t

ORIGIN

Query Match 1.28; Score 20.8; DB 6; Length 46;

Best Local Similarity 78.1%; Pred. No. 3.9e+06;

Matches 25; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Oy 500 TCCCTGAGGGCTACCTGGAGAGCTGACCTC 531

||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 37 TCCCGAGGGCTACCTGCAGAGGCCACCATC 6

Search completed: March 4, 2003, 00:06:28

Job time : 4794 secs

GenCore version 5.1.4.p5.4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 3, 2003, 19:29:35 ; Search time 427 Seconds  
(without alignments)  
9203.131 Million cell updates/sec

Title: US-10-017-621-3

Perfect score: 1745

Sequence: 1 tggagcagctaaagatg.....gttcaactgcccactgtgcc 1745

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 112599159 residues

Total number of hits satisfying chosen parameters: 2166140

Minimum DB seq length: 0

Maximum DB seq length: 50

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N\_Geneseq\_101002:\*

- 1: /SID2/gcgdata/geneseq/geneseq-emb1/NA1980.DAT:\*
- 2: /SID2/gcgdata/geneseq/geneseq-emb1/NA1981.DAT:\*
- 3: /SID2/gcgdata/geneseq/geneseq-emb1/NA1982.DAT:\*
- 4: /SID2/gcgdata/geneseq/geneseq-emb1/NA1983.DAT:\*
- 5: /SID2/gcgdata/geneseq/geneseq-emb1/NA1984.DAT:\*
- 6: /SID2/gcgdata/geneseq/geneseq-emb1/NA1985.DAT:\*
- 7: /SID2/gcgdata/geneseq/geneseq-emb1/NA1986.DAT:\*
- 8: /SID2/gcgdata/geneseq/geneseq-emb1/NA1987.DAT:\*
- 9: /SID2/gcgdata/geneseq/geneseq-emb1/NA1988.DAT:\*
- 10: /SID2/gcgdata/geneseq/geneseq-emb1/NA1989.DAT:\*
- 11: /SID2/gcgdata/geneseq/geneseq-emb1/NA1990.DAT:\*
- 12: /SID2/gcgdata/geneseq/geneseq-emb1/NA1991.DAT:\*
- 13: /SID2/gcgdata/geneseq/geneseq-emb1/NA1992.DAT:\*
- 14: /SID2/gcgdata/geneseq/geneseq-emb1/NA1993.DAT:\*
- 15: /SID2/gcgdata/geneseq/geneseq-emb1/NA1994.DAT:\*
- 16: /SID2/gcgdata/geneseq/geneseq-emb1/NA1995.DAT:\*
- 17: /SID2/gcgdata/geneseq/geneseq-emb1/NA1996.DAT:\*
- 18: /SID2/gcgdata/geneseq/geneseq-emb1/NA1997.DAT:\*
- 19: /SID2/gcgdata/geneseq/geneseq-emb1/NA1998.DAT:\*
- 20: /SID2/gcgdata/geneseq/geneseq-emb1/NA1999.DAT:\*
- 21: /SID2/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:\*
- 22: /SID2/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:\*
- 23: /SID2/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:\*
- 24: /SID2/gcgdata/geneseq/geneseq-emb1/NA2002.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	24.8	1.4	50	22	Human SNP oligonuc
C 2	22.4	1.3	33	24	Human CDK5 related
C 3	22.4	1.3	33	24	Human CDK5 related
C 4	22	1.3	31	22	Human single nucle
C 5	21.4	1.2	31	22	Human single nucle
C 6	21.4	1.2	31	22	Human single nucle
C 7	21.4	1.2	42	22	PCR primer SR53 fo
C 8	21.4	1.2	42	24	ifnc coding region
C 9	21.4	1.2	46	15	Human H1 histone 9

10	21.4	1.2	45	18	AAT63868	Human H1 histone 9
11	21.4	1.2	46	20	AA17156	Test sequence from
12	21.4	1.2	46	24	ABK84647	DNA binding molecu
13	21	1.2	21	22	AAH62195	PCTAIRE-1 polymorp
14	21	1.2	21	22	AAH70659	Human cervical can
15	21	1.2	50	22	AA134286	Human SNP oligonuc
16	20.8	1.2	46	20	AA22932	DE19736591 primer
17	20.6	1.2	45	16	AAT07598	RT-PCR primer/prob
18	20.6	1.2	45	16	AAT00670	Primer 143 for hum
19	20.6	1.2	45	22	AAT27875	Human SNP oligonuc
20	20.4	1.2	32	24	AAT72077	Xcds1 degenerate p
21	20.4	1.2	48	14	AQ50230	HIV pol. INS mutag
22	20.2	1.2	36	19	AAV46356	PCR primer for ser
23	20.2	1.2	40	16	AAQ76190	Human 4-1BB 3' PCR
24	20.2	1.2	40	18	AAT91033	Human cytokine rec
25	20.2	1.2	40	24	ABL54053	Human proton-adeno
26	20.2	1.2	41	24	ABK48669	Human proton-adeno
27	20.2	1.2	41	24	ABK48670	Human prostate spe
28	20.2	1.2	45	22	AAD17287	Human A2b adenosin
29	20.2	1.2	49	20	AAZ31379	MUSIGHAEL Mouse Ig
30	20.2	1.2	50	19	AAV59127	Reverse PCR primer
31	20.2	1.2	50	22	AA543538	Corneddesmosin sin
32	20	1.1	41	19	AAV37843	CD4+ human T-lymph
33	20	1.1	41	24	AAU43619	Human large proteol
34	20	1.1	46	21	AAZ91309	IL-6R and IL-6 fus
35	20	1.1	47	21	AAZ67885	Human map-related
36	20	1.1	50	22	AA129783	Human SNP oligonuc
37	20	1.1	50	22	AA134645	Human SNP oligonuc
38	19.8	1.1	46	21	AA82589	Hammerhead ribozym
39	19.8	1.1	48	18	AAT76056	Human A2b adenosin
40	19.8	1.1	48	20	AA53859	Human adenosine A2
41	19.8	1.1	48	21	AA119424	Human adenosine A2
42	19.8	1.1	48	21	AA333302	Low adenosine anti
43	19.8	1.1	48	21	AAAO3704	Human adenosine A1
44	19.8	1.1	48	22	AAH74232	Nucleotide sequenc
45	19.6	1.1	31	22	AA130029	Human single nucle

#### ALIGNMENTS

RESULT 1

AA134335/C

ID AAL34335 standard; DNA; 50 BP.

XX AAL34335;

AC AAL34335;

DT 24-JAN-2002 (first entry)

DE Human SNP oligonucleotide #7343.

XX Immunosuppressive; immunostimulatory; antiinflammatory; cytostatic;  
 KW neuroprotective; antimicrobial; gene therapy; vaccine; amylase; cancer;  
 KW ankyloid protein; angiotensin; apoptosis related protein; cadherin;  
 KW cyclin; polymerase; oncogene; histone; kinase; colony stimulating factor;  
 KW complement related protein; cytochrome; kinesin; cytokine; interferon;  
 KW interleukin; G-protein coupled receptor; thioesterase; inflammation;  
 KW multifactorial disease; autoimmune disease; infection;  
 KW nervous system disease; ss.

OS Homo sapiens.

PN WO200147944-A2.

XX 05-JUL-2001.

XX 28-DEC-2000; 2000WO-US35498.

XX 28-DEC-1999; 99US-0173419.

XX 27-DEC-2000; 2000US-0173419.

XX (CURA-) CURAGEN CORP.

PI Shimkets RA, Leach M;  
 DR WPI; 2001-465210/50.  
 XX Polymorphic nucleic acids encoding e.g. amylases, cyclins, polymerases,  
 PT oncogenes and histones, useful for diagnosing and treating, e.g.  
 PT cancer, autoimmune diseases and infections  
 XX  
 PS Claim 1; Page 3563; 4143pp; English.  
 XX  
 CC The present invention relates to oligonucleotides encoding polymorphic  
 CC variants of proteins related to amylases, amyloid proteins, angiotensin,  
 CC apoptosis related proteins, cadherin, cyclin, polymerase, oncogenes,  
 CC histones, kinases, colony stimulating factors, complement related  
 CC proteins, cytochromes, kinesins, cytokines, interferons, interleukins,  
 CC G-protein coupled receptors and thioesterases. The present sequence is  
 CC one such oligonucleotide. The oligonucleotides and the peptides encoded  
 CC by them may be used in the prevention, diagnosis and treatment of  
 CC diseases associated with inappropriate expression of the proteins listed  
 CC above. Disorders that may be prevented, diagnosed and/or treated include  
 CC multifactorial diseases with a genetic component, such as autoimmune  
 CC diseases (e.g. rheumatoid arthritis, multiple sclerosis, diabetes,  
 CC systemic lupus erythematosus and Grave's disease), inflammation, cancer  
 CC (e.g. cancers of the bladder, brain, breast, colon and kidney,  
 CC leukaemia), diseases of the nervous system and an infection of pathogenic  
 CC organisms.  
 XX  
 SQ Sequence 50 BP; 7 A; 13 C; 14 G; 16 T; 0 Other;  
 Query Match 1.4%; Score 24.8; DB 22; Length 50;  
 Best Local Similarity 72.7%; Pred. No. 7.8e+03;  
 Matches 32; Conservative 0; Mismatches 12; Indels 0; Gaps 0;  
 OY 1619 CAGACCGAGGCCCGAGCAGCGCGGTGGAGGATGCCACACC 1662  
 DB 45 CAGACCGAGGCCCGAGCAGCTCATCTGCGAGAAATTATGACACC 2  
 RESULT 2  
 ABA04099  
 ID ABA04099 standard; DNA; 33 BP.  
 XX  
 AC ABA04099;  
 XX  
 DT 21-FEB-2002 (first entry)  
 XX  
 DE Human Cdk5 related PCR primer SEQ ID NO:18.  
 XX  
 KW Human; beta-amyloid; cyclin-dependent kinase inhibitor; nerve cell;  
 KW amyloid precursor protein; APP; Cdk5; PCR primer; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PN W0200182967-A1.  
 XX  
 PD 08-NOV-2001.  
 XX  
 PP 25-APR-2001; 2001WO-JP03555.  
 XX  
 PR 28-APR-2000; 2000JP-0131037.  
 XX  
 PA (YAMA) YAMANOUCHI PHARM CO LTD.  
 PA (SUZU) SUZUKI T.  
 XX  
 PI Suzuki T, Watanabe T, Kawabata S, Hachiya S;  
 DR WPI; 2002-026209/03.  
 XX  
 CC Medicinal compositions for the treatment of dementia and Alzheimer's  
 PT disease, comprise compounds that suppress beta amyloid production  
 XX  
 PS Example 6; Page 23; 62pp; Japanese.

CC The present invention describes medicinal compositions (I) inhibiting  
 CC beta-amyloid production comprising an active component a substance that  
 CC inhibits the activity of cyclin-dependent kinase (CDK). Also described  
 CC are: (1) a method for screening compounds for their ability to inhibit  
 CC the production of beta-amyloid by contacting with beta-amyloid producing  
 CC cells; and (2) screening kits. (I) have neurotropic and neuroprotective  
 CC activities. (I) suppress the phosphorylation of amyloid precursor protein  
 CC (APP) which is an essential step in the production of beta-amyloid. (I)  
 CC can be used in the treatment and prevention of neurodegenerative diseases  
 CC such as dementia and Alzheimer's disease. The present sequence represents  
 CC a PCR primer which is used in the exemplification of the present  
 XX invention.  
 SQ Sequence 33 BP; 6 A; 6 C; 11 G; 10 T; 0 Other;  
 Query Match 1.3%; Score 22.4; DB 24; Length 33;  
 Best Local Similarity 81.2%; Pred. No. 2.8e+04;  
 Matches 26; Conservative 0; Mismatches 6; Indels 0; Gaps 0;  
 OY 1018 GAGCTCAAGCTGGCTGACCTTTGGCTGGCCCG 1049  
 DB 2 GAGCTGAATTTGGCTAATTTTGGCTGGCTCG 33  
 RESULT 3  
 ABA04100/c  
 ID ABA04100 standard; DNA; 33 BP.  
 XX  
 AC ABA04100;  
 XX  
 DT 21-FEB-2002 (first entry)  
 XX  
 DE Human Cdk5 related PCR primer SEQ ID NO:19.  
 XX  
 KW Human; beta-amyloid; cyclin-dependent kinase inhibitor; nerve cell;  
 KW amyloid precursor protein; APP; Cdk5; PCR primer; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PN W0200182967-A1.  
 XX  
 PD 08-NOV-2001.  
 XX  
 PP 25-APR-2001; 2001WO-JP03555.  
 XX  
 PR 28-APR-2000; 2000JP-0131037.  
 XX  
 PA (YAMA) YAMANOUCHI PHARM CO LTD.  
 PA (SUZU) SUZUKI T.  
 XX  
 PI Suzuki T, Watanabe T, Kawabata S, Hachiya S;  
 DR WPI; 2002-026209/03.  
 XX  
 CC Medicinal compositions for the treatment of dementia and Alzheimer's  
 PT disease, comprise compounds that suppress beta amyloid production  
 XX  
 PS Example 6; Page 23; 62pp; Japanese.  
 XX  
 CC The present invention describes medicinal compositions (I) inhibiting  
 CC beta-amyloid production comprising an active component a substance that  
 CC inhibits the activity of cyclin-dependent kinase (CDK). Also described  
 CC are: (1) a method for screening compounds for their ability to inhibit  
 CC the production of beta-amyloid by contacting with beta-amyloid producing  
 CC cells; and (2) screening kits. (I) have neurotropic and neuroprotective  
 CC activities. (I) suppress the phosphorylation of amyloid precursor protein  
 CC (APP) which is an essential step in the production of beta-amyloid. (I)  
 CC can be used in the treatment and prevention of neurodegenerative diseases  
 CC such as dementia and Alzheimer's disease. The present sequence represents  
 CC a PCR primer which is used in the exemplification of the present  
 XX invention.  
 SQ Sequence 33 BP; 10 A; 11 C; 6 G; 6 T; 0 other;



```

DT 18-OCT-2001 (first entry)
XX Human single nucleotide polymorphism (SNP) PCTAIRE3 1.
DB Human; resequence; genotype; disease; forensic; paternity testing;
XX single nucleotide polymorphism; SNP; ss.
RW Homo sapiens.
XX
XX Key Location/Qualifiers
FH Variation replace(16.C)
FT /*tag= a
FT /standard_name= "single nucleotide polymorphism"
XX
PN WO200156800-A2.
XX 13-SEP-2001.
XX
XX 07-MAR-2001; 2001WO-US07268.
XX
XX 07-MAR-2000; 2000US-0187510.
XX 22-MAY-2000; 2000US-0206129.
XX (WHED ) WHITEHEAD INST BIOMEDICAL RES.
XX Cargill M, Ireland JS, Lander ES;
XX NPI; 2001-522952/57.
XX
XX Nucleic acid molecules from the human genome which include polymorphic
XX sites, useful in methods for predicting the presence, absence or
XX severity of a particular phenotype or disorder (e.g. diabetes)
XX associated with a particular genotype -
XX
XX Claim 1; Page 34; 145pp; English.
XX
XX The invention relates to the identification of nucleic acid molecules
XX (AA129513-AA13114) from the human genome which include polymorphic sites
XX which can predispose individuals to disease. Various genes from a number
XX of individuals were resequenced and single nucleotide polymorphisms
XX (SNPs) in these genes discovered. The method is useful for predicting the
XX presence, absence or severity of a particular phenotype or disorder (e.g.
XX diabetes) associated with a particular genotype. The nucleic acids
XX containing the polymorphic sites may be useful in forensics and paternity
XX testing.
XX
XX Sequence 31 BP; 6 A; 9 C; 8 G; 8 T; 0 other;
SQ
Query Match 1.2%; Score 21.4; DB 22; Length 31;
Best Local Similarity 80.6%; Pred. No. 4.9e+04;
Matches 25; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
Qy 577 GTCAGCCTATCTGAGATTGCTTGGGAAC 607
Db 1 GCCTCCCTGTCAGACATTGCTTGGGAAC 31
RESULT 7
AAH22523/c
ID AAH22523 standard; DNA; 42 BP.
XX
XX AAH22523;
AC
XX
XX 22-AUG-2001 (first entry)
DT
XX
XX PCR primer SR53 for amplifying a 1fng coding region.
XX
XX Transplastome; plastome; plastid; chloroplast; transgene; plant;
XX 1fng; PCR primer; ss.
XX
XX Synthetic.
XX
XX WO200142441-A2.
XX

```

```

XX
PD 14-JUN-2001.
XX
XX 08-DEC-2000; 2000WO-EP12446.
XX
XX 08-DEC-1999; 99GB-0029075.
XX 14-JUL-2000; 2000GB-0017369.
XX
XX (ITGE-) INT CENT GENETIC ENG & BIOTECHNOLOGY.
XX Reddy S, Sadhu L, Shukla V, Ferraiolo G;
XX WPI; 2001-381671/40.
XX
XX Obtaining a stable transplastome for producing a transplastomic cell,
XX plant or seed, comprises transforming a recipient plastome with a
XX polynucleotide comprising a 5' and 3' sequence homologous to the
XX recipient -
XX
XX Example 12; Page 127; 128pp; English.
XX
XX The invention relates to a method of obtaining a stable transplastome,
XX by transforming a recipient plastome (RP) with a polynucleotide having a
XX 5' sequence homologous to a region of RP, and joined to it a sequence
XX heterologous to RP comprising a coding region operably linked to
XX regulatory region capable of securing expression of coding region in the
XX plastid and joined to it, and a 3' sequence homologous to a region of RP.
XX The method is useful for obtaining a transplastomic plastid, by
XX transforming a plastome within a plastid such as proplastid, amyloplast,
XX chromoplast, etioplast or leucoplast, preferably chloroplast. The method
XX is useful for obtaining a transplastomically expressed protein. The
XX method provides high, uniform, reliable expression of transgenes in
XX plants, with stable inheritance of the trait by avoiding the potential
XX for the dangerous spread of transgenes to the ecosystem. The present
XX sequence represents a PCR primer for amplifying a 1fng coding region,
XX used in generating expression vectors for 1fng in chloroplasts.
XX
XX Sequence 42 BP; 15 A; 10 C; 11 G; 6 T; 0 other;
SQ
Query Match 1.2%; Score 21.4; DB 22; Length 42;
Best Local Similarity 80.6%; Pred. No. 5.6e+04;
Matches 25; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
Qy 270 ACGTGGCTCTCTGGGAACTTCGTTCTGCA 300
Db 35 ACGTACGGGTCTCTGGGAACTTCGTTCTGCA 5
RESULT 8
AAD29563/c
ID AAD29563 standard; DNA; 42 BP.
XX
XX AAD29563;
AC
XX
XX 07-MAY-2002 (first entry)
DT
XX
XX 1fng coding region DNA amplifying PCR primer, SR53.
DE
XX
XX Transgenic plant; transplastomic plant; medicament; PCR primer; ss.
XX
XX Unidentified.
XX
XX WO200206497-A2.
XX
XX 24-JAN-2002.
XX
XX 13-JUL-2001; 2001WO-EP08132.
XX
XX 14-JUL-2000; 2000GB-0017397.
XX
XX (ITGE-) INT CENT GENETIC ENG & BIOTECHNOLOGY.
XX Reddy VS, Sadhu L;
XX

```

XX DR WPI; 2002-171810/22.

XX PT Producing a protein of interest, e.g., a pharmaceutically active

XX PT protein, comprises expressing a polynucleotide fusion construct in a

XX PT plasmid and producing a fusion protein comprising the protein of

XX PT interest

XX PS Example 1; Page 90; 92pp; English.

XX CC The patent discloses a method of producing a protein of interest which

XX CC involves expressing a polynucleotide fusion construct in a plasmid to

XX CC produce a fusion protein comprising the protein of interest where the

XX CC construct comprises a polynucleotide coding sequence of the protein of

XX CC interest operably linked to a polynucleotide coding sequence of a fusion

XX CC protein partner. The methods of the invention are useful for producing a

XX CC protein of interest which comprises a human protein or its biologically

XX CC active variant or fragment, a pharmaceutically active protein, an IFN-

XX CC gamma or its biologically active variant or fragment, a human IFN-

XX CC for the production of transgenic plants. Methods of the invention are

XX CC also useful for the generation of transplasmidic plant cells, plants

XX CC and seeds. The protein of interest obtained by the methods of the

XX CC invention is useful for the manufacture of a medicament for treating

XX CC a disease condition. The present DNA sequence is a PCR primer which

XX CC is used for amplifying the coding region. This primer is used in

XX CC the exemplification of the invention to generate pcuSIRNG expression

XX CC vector for ifng in chloroplasts.

XX SQ Sequence 42 BP; 15 A; 10 C; 11 G; 6 T; 0 other;

Query Match 1.2%; Score 21.4; DB 24; Length 42;

Best Local Similarity 80.6%; Pred. No. 5.6e+04;

Matches 25; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 270 ACCTGCTCTCTCTGGGAACTTCGTCGTCA 300

DB ||||| ||||| ||||| ||||| |||||

35 ACGTACGGGTCTGGCGACCTTCGATCTCA 5

RESULT 9

AAQ69406

ID AAQ69406 standard; DNA; 46 BP.

XX AC AAQ69406;

XX DT 27-FEB-1995 (first entry)

XX DE Human H1 histone gene FNC16, target region.

XX OS DNA protein-binding assay; test sequence; screening sequence;

XX KW promoter; target; TATA box; Herpes Simplex Virus; HSV;

XX KW origin of replication; UL9; transcription factor; TFIID; ds.

XX OS Synthetic.

XX PN W09414980-A.

XX PD 07-JUL-1994.

XX PF 20-DEC-1993; 93US-0512388.

XX PR 23-DEC-1992; 92US-0996783.

XX PR 17-SEP-1993; 93US-0123936.

XX PA (GENE-) GENELABS TECHNOLOGIES INC.

XX PI Andrews BM, Cantor CR, Edwards CA, Fry KE, Turin LM;

XX WPI; 1994-234711/28.

XX PT Sequence-directed DNA-binding molecules - useful in

XX PT pharmaceuticals and as molecular reagents

XX PS Claim 28; Page 290; 587pp; English.

XX CC A DNA protein-binding assay is provided, useful for screening

XX CC libraries of synthetic or biological cpds. for their ability

XX CC to bind DNA test sequences. The assay is versatile in that any

XX CC number of test sequences can be tested by placing the test sequence

XX CC adjacent to a defined protein-binding screening sequence. Binding

XX CC of mols. to these test sequences changes the binding characteristics

XX CC of the protein mol. to its cognate binding sequence. When such a mol.

XX CC binds the test sequence, the equilibrium of the DNA:protein complexes

XX CC is disturbed, generating changes in the concentration of free DNA probe.

XX CC One application of this method is to eucaryotic general transcription

XX CC factors (e.g. TFIID), where the target region is typically selected

XX CC from DNA sequences adjacent to the binding site for the eucaryotic

XX CC transcription factor. Numerous exemplary test sequences are given:

XX CC the sequences in AAQ69251-731 and AAQ69850 correspond to promoter

XX CC targets (typically, TATA box-contg. sites) for human genes and the

XX CC sequences in AAQ69732-849 correspond to promoter targets for viral genes.

XX CC The test sequences may also be randomly generated. DNA:protein

XX CC interaction may be used for screening purposes, e.g. the Herpes Simplex

XX CC Virus (HSV) origin of replication and UL9 (see AAQ69851-52, AAQ69865 and

XX CC AAQ69891).

XX SQ Sequence 46 BP; 9 A; 14 C; 16 G; 7 T; 0 other;

Query Match 1.2%; Score 21.4; DB 15; Length 46;

Best Local Similarity 71.8%; Pred. No. 5.8e+04;

Matches 28; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

OY 1641 GCGGCTGAGGATGCCACACCCCTCAGGCGGACGCC 1679

DB ||||| ||||| ||||| ||||| ||||| ||||| |||||

2 GCGGTGGATTGGACGCTCCACCAATCAACAGGCGGCGCC 40

RESULT 10

AAQ63868

ID AAT63868 standard; DNA; 46 BP.

XX AC AAT63868;

XX DT 14-MAR-1997 (first entry)

XX DE Human H1 histone gene FNC16 gene TFIID binding site.

XX KW Duplex DNA; target region; binding characteristic; DNA binding protein;

XX KW TFIID; transcription factor; binding site; inhibition; enhance;

XX KW cancer; inherited genetic disorder; alpha-D-galactosidase A; ds.

XX OS Homo sapiens.

XX PN US5578444-A.

XX PD 26-NOV-1996.

XX PF 27-JUN-1991; 91US-0723618.

XX PR 20-DEC-1993; 93US-0171389.

XX PR 27-JUN-1991; 91US-0723618.

XX PR 23-DEC-1992; 92US-0996783.

XX PR 17-SEP-1993; 93US-0123936.

XX PA (GENE-) GENELABS TECHNOLOGIES INC.

XX PI Andrews BM, Cantor CR, Edwards CA, Fry KE, Turin LM;

XX WPI; 1997-020402/02.

XX PT Altering binding characteristics of DNA binding proteins to duplex

XX PT DNA - by attaching specific small cpd. to target region close to the

XX PT protein's binding site, useful in treatment of viral disease, cancer

XX PT etc





XX PS Example 15; SEQ ID No 156; 98pp; English.

XX CC The invention relates to a method of decreasing transcriptional activity

XX CC in a duplex deoxyribonucleic acid (DNA) template (T1) comprising

XX CC contacting (T1) with a binding agent comprising at least one small duplex

XX CC DNA-binding molecule (T2) coupled to at least one other small duplex-

XX CC binding molecule that binds to a non-overlapping region of target

XX CC sequence (T5). The method is useful for inhibiting transcription of a

XX CC range of disease-related genes for treating infections (by viruses,

XX CC including human immunodeficiency virus, bacteria, fungi, protozoa

XX CC and parasites), cancer, cardiovascular, respiratory, gastrointestinal,

XX CC endocrine/metabolic, rheumatic/immunological, hematological,

XX CC neurological, psychiatric, dermatological, ophthalmological,

XX CC musculo-skeletal, genetic or urogenital disorders. The method provides

XX CC sequence-specific inhibition of transcription of pathological genes

XX CC without affecting transcription of cellular genes regulated by the same

XX CC transcription factor, and can be applied to regulation of any gene.

XX CC ABK87492-ABK83155 represent DNA binding molecule test sequences used in

XX CC the method of the invention.

XX SQ Sequence 46 BP; 9 A; 14 C; 16 G; 7 T; 0 other;

Query Match 1.2%; Score 21.4; DB 24; Length 46;

Best Local Similarity 71.8%; Pred. No. 5.8e+04;

Matches 28; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 1641 GCGGCTGCAGGATCCACACCCCTCACAGGGCAGGCC 1679

DB 2 GCGGTGGATTGGCGCTCCACCAATCACAGGGCAGGCC 40

RESULT 13

AAH62195

ID AAH62195 standard; DNA; 21 BP.

XX AC AAH62195;

XX DT 12-SEP-2001 (first entry)

XX DE PCTAIRE-1 polymorphism containing DNA fragment #96.

XX KW Single nucleotide polymorphism; SNP; human; cancer; inflammation;

XX KW heart disease; paternity testing; forensic science; ds.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers

XX FT Variation replace(11,G)

XX FT /tag= a

XX FT /standard\_name= "single nucleotide polymorphism"

XX PN WO200138576-A2.

XX PD 31-MAY-2001.

XX PF 17-NOV-2000; 2000WO-US31639.

XX PR 24-NOV-1999; 99US-0167334.

XX PA (WHEED ) WHITEHEAD INST BIOMEDICAL RES.

XX PI Cargill M, Ireland JS, Lander ES;

XX DR WPI; 2001-367705/38.

XX PT New nucleic acid segments of the human genome, particularly from genes

XX PT including polymorphic sites, for phenotype correlation, forensics,

XX PT paternity testing, medicine and genetic analysis

XX PS Claim 1; Page 37; 80pp; English.

XX SQ DNA sequences AAH62100 - AAH62688 represent segments of human genes which

CC contain single nucleotide polymorphisms (SNPs). A method is included in

CC the invention for analysing a nucleic acid sample, which consists of

CC determining the base occupying any one of the polymorphic sites given in

CC the SNP containing sequences. The nucleotide sequences can be used in the

CC diagnosis or monitoring of diseases, such as cancer, inflammation, heart

CC diseases, diseases of the cardiovascular system, and infection by

CC microorganisms. The oligonucleotides are also useful in the manufacture

CC of a medicament for the treatment or prophylaxis of the disease, and as

CC a pharmaceutical SNP containing oligonucleotides are useful in

CC applications such as phenotype correlation, forensics, paternity testing,

CC medicine and genetic analysis.

XX SQ Sequence 21 BP; 9 A; 4 C; 6 G; 2 T; 0 other;

Query Match 1.2%; Score 21; DB 22; Length 21;

Best Local Similarity 100.0%; Pred. No. 5.4e+04;

Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 702 CAAGGAGATCAGACTGGAAACA 722

DB 1 CAAGGAGATCAGACTGGAAACA 21

RESULT 14

AAH70659/c

ID AAH70659 standard; cDNA; 46 BP.

XX AC AAH70659;

XX DT 19-SEP-2001 (first entry)

XX DE Human cervical cancer marker nucleic acid 1933.

XX KW Cervical cancer; cytostatic; pre-malignant condition; gene therapy; ss.

XX OS Homo sapiens.

XX PN WO200142467-A2.

XX PD 14-JUN-2001.

XX PF 08-DEC-2000; 2000WO-US33312.

XX PR 08-DEC-1999; 99US-0169681.

XX PR 21-DEC-1999; 99US-0171350.

XX PR 14-MAR-2000; 2000US-0189315.

XX PR 12-MAY-2000; 2000US-0203791.

XX PR 09-JUN-2000; 2000US-0210600.

XX PR 21-JUL-2000; 2000US-0220114.

XX PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

XX PI Schlegel R, Deeds J, Berger A, Zhao X;

XX DR WPI; 2001-375006/39.

XX PT New isolated nucleic acid for diagnosing and treating cervical cancer

XX PT and for assessing and detecting compounds for treating the cancer -

XX PS Claim 1; Page 415; 1051pp; English.

XX CC The invention relates to novel genes (AAH68727-AAH73383) associated with

XX CC cervical cancer with cytostatic activity. The nucleic acids and encoded

XX CC polypeptides are useful: to assess if a patient is afflicted with

XX CC cervical cancer or has a pre-malignant condition; to monitor the

XX CC progression of cervical cancer or a premalignant condition in a patient;

XX CC and to select and/or assess the efficacy of a compound or therapy for

XX CC inhibiting cervical cancer in a patient. The nucleic acids may also be

XX CC useful for gene therapy.

XX SQ Sequence 46 BP; 7 A; 13 C; 18 G; 6 T; 2 other;

Query Match 1.2%; Score 21; DB 22; Length 46;

Best Local Similarity 71.1%; Pred. No. 7.3e+04;  
Matches 27; Conservative 0; Mismatches 11

**OY** 550 AAGCGCCCTCAGCGCGCGCTCCGTCGTCTCAGCCCTATC 587  
||||| ||| ||| ||||| ||| | ||| |||||  
**D6** 41 AAGCGTCTCTCAGCGCGGCNCCGGGAGTGCTCTCATC 4

## RESULT 15

AAAL34286/C

ID AAL34286 standard; DNA: 50 bp.

AC AAL34286:

24-JAN-2002 (first entry)

Human SNP oligonucleotide #7494.

Immunosuppressive; immunostimulatory; antiinflammatory; cytostatic;  
 neuroprotective; antiproliferative; gene therapy; vaccine; cancer;  
 amyloid protein; antigen; antiproliferative; antiproliferative;  
 cyclin; polymerase; oncogene; growth related protein; cadherin;  
 complement related protein; cytochrome; kinase; cytokine; tumour factor;  
 interleukin; G-protein coupled receptor; thioesterase; inflammation;  
 multifactorial disease; autoimmune disease; infection;  
 nervous system disease; ss.

OS Homo sapiens.

PN WO200147944-A2.

05-JUL-2001.

28-DEC-2000: 2000WO-US35498.

28-DEC-1999; 99US-0173419.

PR 27-DEC-2000; 2000US-0173419;

PA (CURA-) CURAGEN CORP.

PI Shinkets RA, Leach M;

DR WPI; 2001-465210/50.

Polymorphic nucleic acids encoding e.g. amylases, cyclins, polymerases, oncogenes and histones, useful for diagnosing and treating, e.g. cancer, autoimmune diseases and infections -

Claim 1; Page 3547; 4143pp: English.

The present invention relates to oligonucleotides encoding polymorphic variants of proteins related to amylases, amylid proteins, angiotensin, apoptosis related proteins, cadherin, cyclin, polymerase, oncogenes, histones, kinases, colony stimulating factors, complement related proteins, cytochromes, kinesins, cytokines, interferons, interleukins, G-protein coupled receptors and thioesterases. The present sequence is one such oligonucleotide. The oligonucleotides and the peptides encoded by them may be used in the prevention, diagnosis and treatment of diseases associated with inappropriate expression of the proteins listed above. Disorders that may be prevented, diagnosed and/or treated include multifactorial diseases with a genetic component, such as autoimmune diseases (e.g. rheumatoid arthritis, multiple sclerosis, diabetes, systemic lupus erythematosus and Grave's disease), inflammation, cancer (e.g. cancers of the bladder, brain, breast, colon and kidney, leukemia), diseases of the nervous system and an infection of pathogenic organisms.

Sequence 50 BP; 1 A; 27 C; 11 G; 11 T; 0 other;

Query Match  
1.23: Score 21: DB 22: Length 50:

Best Local Similarity 66.7%; Pred. No. 7.6e+04;

Matches 30; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

**Qy** 208 GAGCAGATAGGCCTGGATCAGACTGCTGCTGCTGCCGGCAGTGAC 252  
||||||| ||| ||||| | ||| | ||||| |||  
**Dd** 50 GAGCAGAGAGCGCGGATGGCGCCGGCGCGGAGGGCAGCGGAC 6

Search completed: March 3, 2003, 22:46:23  
Job time : 431 secs



DB 2 GGACGAATCCACCATGGGTGGCTCCCACTCTCTCTCTCTG 44

## RESULT 2

PCT-US95-13830-5

: Sequence 5, Application PC/TUS9513830

: GENERAL INFORMATION:

: APPLICANT: Genentech, Inc.

: TITLE OF INVENTION: Methods and Kits Using Macrophage Stimulating

: TITLE OF INVENTION: Protein

: NUMBER OF SEQUENCES: 10

: CORRESPONDENCE ADDRESS:

: ADDRESSEE: Genentech, Inc.

: STREET: 460 Point San Bruno Blvd

: CITY: South San Francisco

: STATE: California

: COUNTRY: USA

: ZIP: 94080

: COMPUTER READABLE FORM:

: MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk

: COMPUTER: IBM PC compatible

: OPERATING SYSTEM: PC-DOS/MS-DOS

: SOFTWARE: WinPatIn (Genentech)

: CURRENT APPLICATION DATA:

: APPLICATION NUMBER: PCT/US95/13830

: FILING DATE:

: CLASSIFICATION:

: ATTORNEY/AGENT INFORMATION:

: NAME: Marschang, Diane L.

: REGISTRATION NUMBER: 35,600

: TELECOMMUNICATION INFORMATION:

: TELEPHONE: 415/725-5416

: TELEFAX: 415/952-9881

: TELEX: 910/371-7168

: INFORMATION FOR SEQ ID NO: 5:

: SEQUENCE CHARACTERISTICS:

: LENGTH: 47 base pairs

: TYPE: Nucleic Acid

: STRANDEDNESS: Single

: TOPOLOGY: Linear

PCT-US95-13830-5

Query Match 1.34; Score 22.2; DB 5; Length 47;

Best Local Similarity 69.84; Pred. No. 5.8e+03;

Matches 30; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

OY 241 GGCGGAGTGACCTGGAGAGGCCCCACACGCTGCTCTCTG 283

DB 2 GGACGAATCCACCATGGGTGGCTCCCACTCTCTCTCTG 44

## RESULT 3

US-08-507-426C-9/c

: Sequence 9, Application US/08507426C

: Patent No. 6265634

: GENERAL INFORMATION:

: APPLICANT: Lenexa, Philippe

: TITLE OF INVENTION: POLYRIBOSOME CAPABLE OF CONFERRING ON PLANTS RESISTANCE

: TITLE OF INVENTION: TO VIRUSES AND RESISTANT PLANTS PRODUCING THIS

: TITLE OF INVENTION: POLYRIBOSOME

: FILE REFERENCE: 43944-A-PCT-US

: CURRENT APPLICATION NUMBER: US/08/507,426C

: CURRENT FILING DATE: 1995-10-25

: PRIOR FILING DATE: 1995-10-25

: PRIOR APPLICATION NUMBER: 43944-A-PCT-US

: NUMBER OF SEQ ID NOS: 14

: SOFTWARE: Patent In Ver. 2.1

: SEQ ID NO 9

: LENGTH: 44

: TYPE: DNA

: ORGANISM: Artificial Sequence

: FEATURE:  
: OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
: OTHER INFORMATION: ribozymes and portions thereof  
US-08-507-426C-9

Query Match 1.28; Score 21.8; DB 4; Length 44;

Best Local Similarity 70.74; Pred. No. 7.2e+03;

Matches 29; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

OY 1559 CGTCGATGCTGACTCAGCAGGCGACGCTTTCGGGTGTG 1599

DB 42 CTTTGGTGGACGACTCAGCAGCTCTCTTGGCGCTGCTG 2

## RESULT 4

US-08-171-389-156

: Sequence 156, Application US/08171389

: Patent No. 5578444

: GENERAL INFORMATION:

: APPLICANT: Edwards, Cynthia A.

: APPLICANT: Cantor, Charles R.

: APPLICANT: Andrews, Beth M.

: APPLICANT: Turin, Lisa M.

: APPLICANT: Fiy, Kirk E.

: TITLE OF INVENTION: Sequence-Directed DNA Binding

: TITLE OF INVENTION: Molecules, Compositions and Methods

: NUMBER OF SEQUENCES: 641

: CORRESPONDENCE ADDRESS:

: ADDRESSEE: Genelabs Technologies, Inc.

: STREET: 505 Penobscot Drive

: CITY: Redwood City

: STATE: CA

: COUNTRY: USA

: ZIP: 94063

: COMPUTER READABLE FORM:

: MEDIUM TYPE: Floppy disk

: COMPUTER: IBM PC compatible

: OPERATING SYSTEM: PC-DOS/MS-DOS

: SOFTWARE: Patent In Release #1.0, Version #1.25

: CURRENT APPLICATION DATA:

: APPLICATION NUMBER: US/08/171,389

: FILING DATE:

: CLASSIFICATION: 435

: PRIOR APPLICATION DATA:

: APPLICATION NUMBER: US 08/123,936

: FILING DATE: 17-SEP-1993

: PRIOR APPLICATION DATA:

: APPLICATION NUMBER: US 07/995,783

: FILING DATE: 23-DEC-1992

: PRIOR APPLICATION DATA:

: APPLICATION NUMBER: US 07/723,618

: FILING DATE: 27-JUN-1991

: PRIOR APPLICATION DATA:

: APPLICATION NUMBER: US 08/081,070

: FILING DATE: 22-JUN-1993

: ATTORNEY/AGENT INFORMATION:

: NAME: Friblad, Gary R.

: REGISTRATION NUMBER: 33,875

: REFERENCE/DOCKET NUMBER: 4600-0175/G19P3

: TELECOMMUNICATION INFORMATION:

: TELEPHONE: (415) 324-0880

: TELEFAX: (415) 324-0960

: INFORMATION FOR SEQ ID NO: 156:

: SEQUENCE CHARACTERISTICS:

: LENGTH: 46 base pairs

: TYPE: nucleic acid

: STRANDEDNESS: double

: TOPOLOGY: linear

: MOLECULE TYPE: DNA (genomic)

: HYPOTHETICAL: NO

: ORIGINAL SOURCE:

: INDIVIDUAL ISOLATE: Human H1 histone gene FNCL6

US-08-171-389-156

Query Match 1.2% Score 21.4; DB 1; Length 46;  
 Best Local Similarity 71.8%; Pred. No. 9.3e+03;  
 Matches 28; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 1641 GCGGCTGGAGGATGCCACACCCCTCACAGGGCAGCCCC 1679  
 DB 2 GCGGTGGATTGGACGCTCCACCAATCACAGGGCAGCGCC 40

## RESULT 5

US-08-123-936-156  
 : Sequence 156, Application US/08123936  
 : Patent No. 5726014

## GENERAL INFORMATION:

APPLICANT: Edwards, Cynthia A.  
 APPLICANT: Cantor, Charles R.  
 APPLICANT: Andrews, Beth M.  
 APPLICANT: Turin, Lisa M.  
 TITLE OF INVENTION: Screening Assay for the Detection of  
 : DNA-Binding Molecules  
 : NUMBER OF SEQUENCES: 640  
 : CORRESPONDENCE ADDRESS:  
 : ADDRESSEE: Genelabs Technologies, Inc.  
 : STREET: 505 Penobscot Drive  
 : CITY: Redwood City  
 : STATE: CA  
 : COUNTRY: USA  
 : ZIP: 94063

COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/123,936  
 FILING DATE: 27-JUN-1991  
 CLASSIFICATION: 435

PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 07/996,783  
 FILING DATE: 23-DEC-1992  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 07/723,618  
 FILING DATE: 27-JUN-1991

ATTORNEY/AGENT INFORMATION:  
 NAME: Fabian, Gary R.  
 REGISTRATION NUMBER: 33,875  
 REFERENCE/DOCKET NUMBER: 4600-0075.32/G19P2  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (415) 324-0960  
 TELEFAX: (415) 324-0960  
 INFORMATION FOR SEQ ID NO: 156:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 46 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: double  
 TOPOLOGY: linear  
 MOLECULE TYPE: DNA (genomic)  
 HYPOTHETICAL: NO  
 ORIGINAL SOURCE:  
 INDIVIDUAL ISOLATE: Human H1 histone gene FNC16

US-08-123-936-156  
 Query Match 1.2% Score 21.4; DB 1; Length 46;  
 Best Local Similarity 71.8%; Pred. No. 9.3e+03;  
 Matches 28; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 1641 GCGGCTGGAGGATGCCACACCCCTCACAGGGCAGCCCC 1679  
 DB 2 GCGGTGGATTGGACGCTCCACCAATCACAGGGCAGCGCC 40

## RESULT 6

US-08-475-228A-156  
 : Sequence 156, Application US/08475228A  
 : Patent No. 5869241  
 : GENERAL INFORMATION:  
 : APPLICANT: Edwards, Cynthia A.  
 : APPLICANT: Cantor, Charles R.  
 : APPLICANT: Andrews, Beth M.  
 : APPLICANT: Turin, Lisa M.  
 : APPLICANT: Evi, Kirk E.  
 : TITLE OF INVENTION: Sequence-Directed DNA Binding  
 : MOLECULES, COMPOSITIONS AND METHODS  
 : NUMBER OF SEQUENCES: 664  
 : CORRESPONDENCE ADDRESS:  
 : ADDRESSEE: Genelabs Technologies, Inc.  
 : STREET: 505 Penobscot Drive  
 : CITY: Redwood City  
 : STATE: CA  
 : COUNTRY: USA  
 : ZIP: 94063

COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/475,228A  
 FILING DATE: 06-JUN-1995  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 08/123,936  
 FILING DATE: 17-SEP-1993

PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 07/996,783  
 FILING DATE: 23-DEC-1992  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 07/723,618  
 FILING DATE: 27-JUN-1991

PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 08/081,070  
 FILING DATE: 22-JUN-1993  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Stratford, Carol A.  
 REGISTRATION NUMBER: 34,444  
 REFERENCE/DOCKET NUMBER: 4600-0175.21/G19P3D2  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (415) 324-0960  
 TELEFAX: (415) 324-0960

INFORMATION FOR SEQ ID NO: 156:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 46 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: double  
 TOPOLOGY: linear  
 MOLECULE TYPE: DNA (genomic)  
 HYPOTHETICAL: NO  
 ORIGINAL SOURCE:  
 INDIVIDUAL ISOLATE: Human H1 histone gene FNC16

US-08-475-228A-156  
 Query Match 1.2% Score 21.4; DB 2; Length 46;  
 Best Local Similarity 71.8%; Pred. No. 9.3e+03;  
 Matches 28; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 1641 GCGGCTGGAGGATGCCACACCCCTCACAGGGCAGCCCC 1679  
 DB 2 GCGGTGGATTGGACGCTCCACCAATCACAGGGCAGCGCC 40

RESULT 7  
 US-08-482-080A-156  
 : Sequence 156, Application US/08482080A  
 : Patent No. 6010849  
 : GENERAL INFORMATION:  
 : APPLICANT: Edwards, Cynthia A.

APPLICANT: Turin, Lisa M.  
 APPLICANT: Ery, Kirk E.  
 TITLE OF INVENTION: Sequence-Directed DNA Binding  
 TITLE OF INVENTION: Molecules, Compositions and Methods  
 NUMBER OF SEQUENCES: 664  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: GeneLabs Technologies, Inc.  
 STREET: 505 Penobscot Drive  
 CITY: Redwood City  
 STATE: CA  
 COUNTRY: USA  
 ZIP: 94063  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patentia Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/354,947  
 FILING DATE:  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 08/482,080  
 FILING DATE: 07-JUN-1995  
 APPLICATION NUMBER: US 08/171,389  
 FILING DATE: 20-DEC-1993  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 08/123,936  
 FILING DATE: 17-SEP-1993  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 07/996,783  
 FILING DATE: 23-DEC-1992  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 07/723,618  
 FILING DATE: 27-JUN-1991  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 08/081,070  
 FILING DATE: 22-JUN-1993  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Brady, John F.  
 REGISTRATION NUMBER: 39,118  
 REFERENCE/DOCKET NUMBER: 4600-0175.20/G19P3D1  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (650) 324-0880  
 TELEFAX: (650) 324-0960  
 INFORMATION FOR SEQ ID NO: 156:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 46 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: double  
 TOPOLOGY: linear  
 MOLECULE TYPE: DNA (genomic)  
 HYPOTHETICAL: NO  
 ORIGINAL SOURCE:  
 INDIVIDUAL ISOLATE: Human H1 histone gene FNC16  
 US-09-354-947-156

Query Match 1.28; Score 21.4; DB 4; Length 46;  
Best Local Similarity 71.88; Pred. No. 9.3e+03;  
Matches 28; Conservative 0; Mismatches 11; Indels 0; Caps 0;

QY	1641	GCGGCTGGAGGGATGCCACACCCCTCACAGGGCAGCC	1679
DB	2	GCGGTGATGGAGCTCCACCAATCACAGGGCAGCC	40

```

RESULT 9
PCT-US93-12388-156
; Sequence 156, Application PC/TUS9312388
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: Sequence-Directed DNA Binding
; TITLE OF INVENTION: Molecules, Compositions and Methods
; NUMBER OF SEQUENCES: 641

```

```

CORRESPONDENCE ADDRESS:
ADDRESSEE: Genelabs Technologies, Inc.
STREET: 505 Pencabscot Drive
CITY: Redwood City
STATE: CA
COUNTRY: USA
ZIP: 94063
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US93/12388
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/123,936
FILING DATE: 17-SEP-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/996,783
FILING DATE: 23-DEC-1992
ATTORNEY/AGENT INFORMATION:
NAME: Fabian, Gary R.
REGISTRATION NUMBER: 33,875
REFERENCE/DOCKET NUMBER: 4600-0175.41/G19PCT2
TELEPHONE: (415) 324-0880
TELEFAX: (415) 324-0960
INFORMATION FOR SEQ ID NO: 156:
SEQUENCE CHARACTERISTICS:
LENGTH: 46 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ORIGINAL SOURCE:
INDIVIDUAL ISOLATE: Human H1 histone gene FNC16
PCT-US93-12388-156

Query Match 1.2% Score 21.4; DB 5; Length 46;
Best Local Similarity 71.8%; Pred. No. 9.1e-03;
Matches 28; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

OY 1641 CGCGCTGAGGGATGCCACACCCCTCACAGGCGAGCC 1679
||||| 1 ||| 11 ||| ||||| ||| 11
DB 2 CGGGTGGATTGGACGCTCCACCAATCACAGGCGAGCC 40

RESULT 10
US-08-233-009-33
Sequence 33, Application US/08233009
Patent No. 5646156
GENERAL INFORMATION:
APPLICANT: Jacobson, Marlene A
APPLICANT: Johnson, Robert G
APPLICANT: Salvatore, Christopher A
TITLE OF INVENTION: INHIBITION OF EOSINOPHIL
TITLE OF INVENTION: ACTIVATION THROUGH A3 ADENOSINE RECEPTOR ANTAGONISM
NUMBER OF SEQUENCES: 56
CORRESPONDENCE ADDRESS:
ADDRESSEE: Merck & Co., Inc.
STREET: P.O. Box 2000
CITY: Rahway
STATE: New Jersey
COUNTRY: United States
ZIP: 07065
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25

CORRESPONDENCE ADDRESS:
ADDRESSEE: Genelabs Technologies, Inc.
STREET: 505 Pencabscot Drive
CITY: Redwood City
STATE: CA
COUNTRY: USA
ZIP: 94063
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/233,009
FILING DATE: 23-APR-1994
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Bencen, Gerard H
REGISTRATION NUMBER: 35,746
REFERENCE/DOCKET NUMBER: 19219
TELEPHONE: (908) 594-3901
TELEFAX: (908) 594-4720
INFORMATION FOR SEQ ID NO: 33:
SEQUENCE CHARACTERISTICS:
LENGTH: 45 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: linear
MOLECULE TYPE: CDNA
HYPOTHETICAL: NO
ANTI-SENSE: YES
US-08-233-009-33

Query Match 1.2% Score 20.6; DB 1; Length 45;
Best Local Similarity 67.4%; Pred. No. 1.5e-04;
Matches 29; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

OY 737 CCTGCACCGCCATCGGGAGGTGCTCCTGCTCAAGGACCTCAA 779
|||||| 111 1 11 |||| 1111 11 1
DB 2 CCTGCACCGCCACTGCTTAGCCCCCTCGCCAGGCTCATCCA 44

RESULT 11
US-08-285-936-38/c
Sequence 38, Application US/08285936
Patent No. 5728821
GENERAL INFORMATION:
APPLICANT: Yelton, Dale
APPLICANT: Glaser, Scott
APPLICANT: Huse, William
APPLICANT: Rosok, Mae J.
TITLE OF INVENTION: No. 5728821el Mutant BR96 Antibodies and
TITLE OF INVENTION: Functional Equivalents Reactive With Human Carcinomas
NUMBER OF SEQUENCES: 57
CORRESPONDENCE ADDRESS:
ADDRESSEE: Merchant & Gould
STREET: 11150 Santa Monica Blvd., Suite 400
CITY: Los Angeles
STATE: CA
COUNTRY: USA
ZIP: 90025-3395
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/285,936
FILING DATE: 04-AUG-1994
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Adriano, Sarah B.
REGISTRATION NUMBER: 34,470
REFERENCE/DOCKET NUMBER: 30436.16US01
TELEPHONE: 310-445-1140
TELEFAX: 310-445-9031
INFORMATION FOR SEQ ID NO: 38:
SEQUENCE CHARACTERISTICS:
LENGTH: 45 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
```





MEDIUM TYPE: FLOPPY DISK  
COMPUTER: IBM PC COMPATIBLE  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: WORDPERFECT 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/050,478  
FILING DATE: 26-OCT-1994  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/US93/02908  
FILING DATE: 29-MAR-1993  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/858,747  
FILING DATE: 27-MAR-1992  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: MORRY, MARY J.  
REGISTRATION NUMBER: 34,398  
REFERENCE/DOCKET NUMBER: 2026-4006US1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212)758-4800  
TELEFAX: (212)751-6849  
INFORMATION FOR SEQ ID NO: 47:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 48 BASE PAIRS  
TYPE: NUCLEIC ACID  
STRANDEDNESS: SINGLE  
TOPOLOGY: LINEAR  
US-08-050-478-47

Query Match 1.2%; Score 20.4; DB 2; Length 48;  
Best Local Similarity 71.1%; Pred. No. 1.7e+04;  
Matches 27; Conservative 0; Mismatches 11; Indels 0; Gaps 0;  
Oy 325 GAGATTGTGCACGAGGACTTGAAGATGGGGTCTGATGG 362  
||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 7 GACACGGTGCCTGCTGAAGTTGAAGCGGGGATGGATGG 44

RESULT 15  
US-09-414-117-47  
Sequence 47, Application US/09414117  
Patent No. 6291684  
GENERAL INFORMATION:  
APPLICANT:  
APPLICANT:  
TITLE OF INVENTION: METHOD OF ELIMINATING  
NUMBER OF INVENTION: INHIBITORY/INSTABILITY REGIONS OF mRNA  
NUMBER OF SEQUENCES: 130  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: MORGAN & FINNEGAN  
STREET: 345 PARK AVENUE  
CITY: NEW YORK  
STATE: NEW YORK  
COUNTRY: USA  
ZIP: 10154  
COMPUTER READABLE FORM:  
MEDIUM TYPE: FLOPPY DISK  
COMPUTER: IBM PC COMPATIBLE  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: WORDPERFECT 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/414,117  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/850,049  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/US93/02908

FILING DATE: 29-MAR-1993  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/858,747  
FILING DATE: 27-MAR-1992  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: MORRY, MARY J.  
REGISTRATION NUMBER: 34,398  
REFERENCE/DOCKET NUMBER: 2026-4006US1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212)751-6849  
TELEFAX: (212)751-6849  
INFORMATION FOR SEQ ID NO: 47:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 48 BASE PAIRS  
TYPE: NUCLEIC ACID  
STRANDEDNESS: SINGLE  
TOPOLOGY: LINEAR  
US-09-414-117-47  
Query Match 1.2%; Score 20.4; DB 4; Length 48;  
Best Local Similarity 71.1%; Pred. No. 1.7e+04;  
Matches 27; Conservative 0; Mismatches 11; Indels 0; Gaps 0;  
Oy 325 GAGATTGTGCACGAGGACTTGAAGATGGGGTCTGATGG 362  
||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 7 GACACGGTGCCTGCTGAAGTTGAAGCGGGGATGGATGG 44  
Search completed: March 4, 2003, 00:52:00  
Job time : 80 secs

## SUMMARIES

APPLICANT: Ireland, James S.

PRIOR FILING DATE: 2000-03-07

STRANDED

```
;
; TOPOLOGY: LINEAR
; US-09-943-722-47
;
; Query Match
; Best Local Similarity 1.1%; Score 20.4; DB 9; Length 48;
; Matches 27; Conservative 0; Mismatches 11; Indels 0; Gaps 0;
;
; QY 325 GAGATGTCGACGAGGACTGAGATGGGGTCTGATGG 362
; DB 7 GAGACGGTCCCGTGAAGTTGAAGCCGGGATGGATGG 44
;
; RESULT 6
; US-09-147-142-11
; Sequence 11, Application US/09147142
; Patent No. US20020018749A1
; GENERAL INFORMATION:
; APPLICANT: HUDSON, Peter John
; APPLICANT: KORTT, Alex Andrew
; APPLICANT: IRVING, Robert Alexander
; APPLICANT: ATWELL, John Leslie
; TITLE OF INVENTION: HIGH AVIDITY POLYVALENT AND POLYSPECIFIC REAGENTS
; FILE REFERENCE: 016786/0212
; CURRENT APPLICATION NUMBER: US/09/147.142
; EARLIER FILING DATE: 1999-03-05
; EARLIER APPLICATION NUMBER: PCT/AU98/00212
; EARLIER FILING DATE: 1998-03-26
; EARLIER APPLICATION NUMBER: AU PO 5917
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 11
; LENGTH: 45
; TYPE: DNA
; ORGANISM: Artificial Sequence
;
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: Oligonucleotide used to insert codon between VH
; OTHER INFORMATION: and VL domains of NC10 scFv-0
; US-09-147-142-11
;
; Query Match
; Best Local Similarity 1.1%; Score 20; DB 10; Length 45;
; Matches 29; Conservative 0; Mismatches 15; Indels 0; Gaps 0;
;
; QY 643 GGTACTATGCCACCGTCTACAAAGGCAAAAGCAGGCTCACACA 686
; DB 1 GGGACCACGGTCACCGTCTCCGGTGTGATATCGAGCTCACACA 44
;
; RESULT 7
; US-09-147-142-12/c
; Sequence 12, Application US/09147142
; Patent No. US20020018749A1
; GENERAL INFORMATION:
; APPLICANT: HUDSON, Peter John
; APPLICANT: KORTT, Alex Andrew
; APPLICANT: IRVING, Robert Alexander
; APPLICANT: ATWELL, John Leslie
; TITLE OF INVENTION: HIGH AVIDITY POLYVALENT AND POLYSPECIFIC REAGENTS
; FILE REFERENCE: 016786/0212
; CURRENT APPLICATION NUMBER: US/09/147.142
; EARLIER FILING DATE: 1999-03-05
; EARLIER APPLICATION NUMBER: PCT/AU98/00212
; EARLIER FILING DATE: 1998-03-26
; EARLIER APPLICATION NUMBER: AU PO 5917
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 12
; LENGTH: 45
; TYPE: DNA
; ORGANISM: Artificial Sequence
;
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: Oligonucleotide used to insert codon between VH
; OTHER INFORMATION: and VL domains of NC10 scFv-0
; US-09-147-142-11
;
; Query Match
; Best Local Similarity 1.1%; Score 20; DB 10; Length 45;
; Matches 29; Conservative 0; Mismatches 15; Indels 0; Gaps 0;
;
; QY 643 GGTACTATGCCACCGTCTACAAAGGCAAAAGCAGGCTCACACA 686
; DB 1 GGGACCACGGTCACCGTCTCCGGTGTGATATCGAGCTCACACA 44
;
; RESULT 8
; US-09-263-959-758
; Sequence 758, Application US/09263959
; Patent No. US20020150891A1
; GENERAL INFORMATION:
; APPLICANT: HOOD, Leroy E.
; APPLICANT: KOOP, Ben F.
; APPLICANT: ROVEN, Lee
; TITLE OF INVENTION: DIAGNOSTIC AND THERAPEUTIC COMPOSITIONS AND METHODS WHICH U
; NUMBER OF SEQUENCES: 1279
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Seed and Berry LLP
; CITY: Seattle
; STATE: Washington
; COUNTRY: US
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/263.959
; FILING DATE: 05-MAR-1999
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Mcmasters, David D.
; REGISTRATION NUMBER: 33,963
; REFERENCE/DOCKET NUMBER: 920010.426C2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 758:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 40 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-09-263-959-758
;
; Query Match
; Best Local Similarity 1.1%; Score 19.8; DB 10; Length 40;
; Matches 24; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
;
; QY 1689 CTTCCCTGCTTACTCTCTGCTGCTGCTGCTGCTG 1719
; DB 3 CTTCCCTGCTTCTCTCTCTCTCTCTCTCTCTCTG 33
;
; RESULT 9
; US-09-790-417-235/c
; Sequence 235, Application US/09790417
; Patent No. US20010031470A1
; GENERAL INFORMATION:
; APPLICANT: Shultz, John W
; APPLICANT: Lewis, Martin K.
; APPLICANT: Lieppe, Donna
```

```
; APPLICANT: Mandrekar, Michelle
; APPLICANT: Kephart, Daniel
; APPLICANT: Rhodes, Richard B.
; APPLICANT: Andrews, Christine A.
; APPLICANT: Hartnett, James R.
; APPLICANT: Gu, Trent
; APPLICANT: Olson, Ryan J.
; APPLICANT: Wood, Keith W.
; APPLICANT: Welch, Roy
; TITLE OF INVENTION: Nucleic Acid Detection
; FILE REFERENCE: PRO-103 6868/75528
; CURRENT APPLICATION NUMBER: US/09/790,417
; PRIOR FILING DATE: 2001-03-22
; PRIOR APPLICATION NUMBER: 09/356,972
; PRIOR FILING DATE: 1999-07-21
; PRIOR APPLICATION NUMBER: 09/042,287
; PRIOR FILING DATE: 1998-03-13
; NUMBER OF SEQ ID NOS: 290
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 233
; LENGTH: 42
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:probe for oligo
US-09-790-417-235

Query Match      1.1%; Score 19.6; DB 10; Length 42;
Best Local Similarity 66.7%; Pred. No. 3.5e+04;
Matches 29; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

Oy 759 GTCCCTGCTCAAGGAGCTCAAGACGCGCAACATCGTTACGCT 800
    ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 42 GTACCTGTAATCAACTCACCAGCATATCATCACCAGCT 1

RESULT 10
US-10-073-256-27/c
; Sequence 27, Application US/10073256
; Patent No. US20020120408A1
; GENERAL INFORMATION:
; APPLICANT: Kreiswirth, Barry N
; APPLICANT: Nadich, Steven M
; TITLE OF INVENTION: System and Method for Tracking and Controlling Infections
; FILE REFERENCE: 19124.0002
; CURRENT APPLICATION NUMBER: US/10/073,256
; CURRENT FILING DATE: 2002-02-13
; NUMBER OF SEQ ID NOS: 80
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 27
; LENGTH: 45
; TYPE: DNA
; ORGANISM: Enterococcus faecalis
US-10-073-256-27

Query Match      1.1%; Score 19.4; DB 12; Length 45;
Best Local Similarity 64.4%; Pred. No. 4.1e+04;
Matches 29; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

Oy 1288 ATCTCTCAAGGAGGTCTCAAGACATCACTACCCCAAGTAC 1332
    ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 45 AACCAAGCAACAGTAGTACCAGCGAANGTAGTACCTCAAGTAC 1

RESULT 11
US-10-073-256-35/c
; Sequence 35, Application US/10073256
; Patent No. US20020120408A1
; GENERAL INFORMATION:
; APPLICANT: Kreiswirth, Barry N
; APPLICANT: Nadich, Steven M
; TITLE OF INVENTION: System and Method for Tracking and Controlling Infections
```

```
; FILE REFERENCE: 19124.0002
; CURRENT APPLICATION NUMBER: US/10/073,256
; CURRENT FILING DATE: 2002-02-13
; NUMBER OF SEQ ID NOS: 80
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 35
; LENGTH: 45
; TYPE: DNA
; ORGANISM: Enterococcus faecalis
US-10-073-256-35

Query Match      1.1%; Score 19.4; DB 12; Length 45;
Best Local Similarity 64.4%; Pred. No. 4.1e+04;
Matches 29; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

Oy 1268 ATCTCTCAAGGAGGTCTCAAGACATCACTACCCCAAGTAC 1332
    ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 45 AACCAAGTGAACCAAGTAGTATACAGCGAAGAGCAGCAGCTCAAGTAC 1

RESULT 12
US-09-801-274-517
; Sequence 517, Application US/09801274
; Patent No. US20020032319A1
; GENERAL INFORMATION:
; APPLICANT: Cargill, Michele
; APPLICANT: Ireland, James S.
; APPLICANT: Lander, Eric S.
; TITLE OF INVENTION: HUMAN SINGLE NUCLEOTIDE POLYMORPHISMS
; FILE REFERENCE: 2825.2009-001
; CURRENT APPLICATION NUMBER: US/09/801,274
; CURRENT FILING DATE: 2001-03-07
; PRIOR APPLICATION NUMBER: US 60/187,510
; PRIOR FILING DATE: 2000-03-07
; PRIOR APPLICATION NUMBER: US 60/206,129
; PRIOR FILING DATE: 2000-05-22
; NUMBER OF SEQ ID NOS: 1802
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 517
; LENGTH: 31
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-801-274-517

Query Match      1.1%; Score 19.2; DB 10; Length 31;
Best Local Similarity 80.8%; Pred. No. 3.9e+04;
Matches 21; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Oy 1150 CCACAGCGCTCCCTCTTCCGGGC 1215
    ||||| ||||| ||||| ||||| |||||
Db 5 CCACAGCGCTTCCCTCTTCCGGGC 30

RESULT 13
US-09-263-959-121/c
; Sequence 121, Application US/09263959
; Patent No. US20020150891A1
; GENERAL INFORMATION:
; APPLICANT: Hood, Leroy E.
; APPLICANT: Rowen, Lee
; APPLICANT: Koop, Ben F.
; TITLE OF INVENTION: DIAGNOSTIC AND THERAPEUTIC COMPOSITIONS AND METHODS WHICH U
; NUMBER OF SEQUENCES: 1279
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Seed and Berry LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: US
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
```

```

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/263,959
FILING DATE: 05-MAR-1999
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: McMAsters, David D.
REGISTRATION NUMBER: 33,963
REFERENCE/DOCKET NUMBER: 920010.426C2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 121:
SEQUENCE CHARACTERISTICS:
LENGTH: 46 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-09-263-959-121

```

```

Query Match      1.18; Score 19.2; DB 10; Length 45;
Best Local Similarity 75.04; Pred. No. 4.7e+04;
Matches 24; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
QY 319 GCACCAGAGATTGTGCACGAGGACTTGAAGT 350
      ||||| ||||| ||||| |||||
DB 46 GCACCAGAGTTTCTGCACAGGGAGGGGGAGAT 15

```

```

RESULT 14
US-09-376-940-23/c
    Sequence 23, Application US/09376940A
    Publication No. US20020192813A1
    GENERAL INFORMATION:
    APPLICANT: Conner, Timothy W
    APPLICANT: Santino, Colleen G
    TITLE OF INVENTION: No. US20020192813A1e1 Plant Expression Vectors
    FILE REFERENCE: monocot elements
    CURRENT APPLICATION NUMBER: US/09/376,940A
    CURRENT FILING DATE: 1999-08-18
    EARLIER APPLICATION NUMBER: 60/097150
    EARLIER FILING DATE: 1998-08-19
    NUMBER OF SEQ ID NOS: 63
    SOFTWARE: PatentIn ver. 2.1
    SEQ ID NO 23
    LENGTH: 43
    TYPE: DNA
    ORGANISM: Artificial Sequence
    FEATURE:
    OTHER INFORMATION: Description of Artificial Sequence:synthetic
US-09-376-940-23

```

```

Query Match      1.14; Score 19; DB 9; Length 43;
Best Local Similarity 65.14; pred. No. 5.1e+04;
Matches 28; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 79 GGGGCCCCGGGCTCTGAGTTGCTGGCGCCCCCGCCGATCG 121
Db 43 GTGTCCACCCACCCCTCGATCTCTGGTTCGCCCGCCGATCG 1

```

RESULT 15  
US-09-818-066-32  
; Sequence 32, Application US/09818066  
; Patent No. US2002003207A1  
; GENERAL INFORMATION:  
; APPLICANT: Shuping Tong et al.  
; TITLE OF INVENTION: HEPADNAVIRUS RECEPTOR  
; NUMBER OF SEQUENCES: 75  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fish & Richardson P.C.  
; STREET: 225 Franklin Street

CITY: Boston  
STATE: Massachusetts  
COUNTRY: U.S.A.  
ZIP: 02110-2804  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/818,066  
FILING DATE: 27-Mar-2001  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/683,262  
FILING DATE: 18-JUL-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Fraser, Janis K.  
REGISTRATION NUMBER: 31,819  
REFERENCE/DOCKET NUMBER: 00786/287002  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 542-5070  
TELEFAX: (617) 542-8906  
TELEX: 200154  
INFORMATION FOR SEQ ID NO: 32:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 45 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
SEQUENCE DESCRIPTION: SEQ ID NO: 32:  
US-09-818-066-32

```
Query Match      1.18; Score 19; DB 10; Length 45;
Best Local Similarity 71.4%; Prd. No. 5.2e-04;
Matches 25; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
```

Search completed: March 4, 2003, 00:54:25  
Job time : 134 secs

GenCore version 5.1.4\_P5\_4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OK nucleic - nucleic search, using sw model

Run on: March 3, 2003, 22:34:05 ; Search time 2628 seconds  
(without alignments)  
10753.853 Million cell updates/sec

Title: US-10-017-621-3

Perfect score: 1745

Sequence: 1 tggagcagcgtaaggatg.....gtccactgccactgtcc 1745

Scoring table: IDENTITY NUC

Gapop 10.0, Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 102860

Minimum DB seq length: 0

Maximum DB seq length: 50

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST.\*

1: em\_estb.\*  
2: em\_esthm.\*  
3: em\_estin.\*  
4: em\_estnu.\*  
5: em\_estov.\*  
6: em\_estpl.\*  
7: em\_estro.\*  
8: em\_hci.\*  
9: gb\_esti.\*  
10: gb\_est2.\*  
11: gb\_hic.\*  
12: gb\_est3.\*  
13: gb\_est4.\*  
14: gb\_est5.\*  
15: em\_estfun.\*  
16: em\_estom.\*  
17: gb\_gss.\*  
18: em\_gss\_hum.\*  
19: em\_gss\_inv.\*  
20: em\_gss\_pin.\*  
21: em\_gss\_vrt.\*  
22: em\_gss\_fun.\*  
23: em\_gss\_nam.\*  
24: em\_gss\_hus.\*  
25: em\_gss\_other.\*  
26: em\_gss\_pro.\*  
27: em\_gss\_rtd.\*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	45	2.6	46	14	N78054
2	28	1.6	28	14	R38968
3	23.2	1.3	48	10	AW247978
4	21.4	1.2	50	9	AW247978
5	21.2	1.2	36	17	AZ346286
6	21	1.2	50	9	AU102877

C 7	21	1.2	50	9	AU105237
C 8	21	1.2	46	13	BM397711
C 9	20.8	1.2	46	17	AZ939393
C 10	20.6	1.2	48	17	AZ311362
C 11	20.6	1.2	50	9	AU106960
C 12	20.4	1.2	47	17	AZ331336
C 13	20.2	1.2	45	17	AZ965975
C 14	20.2	1.2	49	9	AZ204601
C 15	20	1.1	40	9	A1475974
C 16	20	1.1	40	17	BH811451
C 17	19.8	1.1	49	14	W39000
C 18	19.4	1.1	50	9	AU104829
C 19	19.2	1.1	44	12	BG422154
C 20	19.2	1.1	49	17	AZ450961
C 21	19.2	1.1	50	9	AU103357
C 22	19.2	1.1	50	9	AU103358
C 23	19.2	1.1	50	9	AU103359
C 24	19.2	1.1	50	9	AU103361
C 25	19.2	1.1	50	9	AU103381
C 26	19.2	1.1	50	9	AU103915
C 27	19.2	1.1	50	9	AU106349
C 28	19.2	1.1	50	14	T74703
C 29	19	1.1	43	9	A1591257
C 30	19	1.1	49	17	AZ423762
C 31	19	1.1	50	9	AU107320
C 32	18.8	1.1	34	9	AA972479
C 33	18.8	1.1	43	17	BH857724
C 34	18.8	1.1	46	17	BH950015
C 35	18.8	1.1	50	9	AU102939
C 36	18.8	1.1	50	9	AU103583
C 37	18.8	1.1	50	9	AU104587
C 38	18.8	1.1	50	9	AU105918
C 39	18.6	1.1	44	17	TA165050
C 40	18.6	1.1	45	9	TA150043
C 41	18.6	1.1	49	17	AZ966392
C 42	18.6	1.1	50	9	AU103382
C 43	18.6	1.1	50	9	AU103553
C 44	18.6	1.1	50	9	AU104162
C 45	18.6	1.1	50	13	BI910989

## ALIGNMENTS

RESULT 1  
N78054  
LOCUS  
DEFINITION  
Yv71905.r1 Soares fetal liver spleen INFEL Homo sapiens cDNA clone IMAGE:248216 5' similar to gb:X66363 SERINE/THREONINE-PROTEIN KINASE PCTAIRE-1 (HUMAN); mRNA sequence.

ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
N78054.1 GI:1240755  
human.

REFERENCE

AUTHORS

REFERENCES

AUTHORS

REFERENCES

AUTHORS

REFERENCES

AUTHORS

REFERENCES

AUTHORS

REFERENCES

AUTHORS

REFERENCES

AUTHORS

REFERENCES

AUTHORS

REFERENCES

AUTHORS

REFERENCES

AUTHORS

REFERENCES

AUTHORS

REFERENCES

AUTHORS

REFERENCES

AUTHORS

46 bp mRNA linear EST 28-JAN-1997  
Yv71905.r1 Soares fetal liver spleen INFEL Homo sapiens cDNA clone IMAGE:248216 5' similar to gb:X66363 SERINE/THREONINE-PROTEIN KINASE PCTAIRE-1 (HUMAN); mRNA sequence.

N78054  
LOCUS  
DEFINITION  
Yv71905.r1 Soares fetal liver spleen INFEL Homo sapiens cDNA clone IMAGE:248216 5' similar to gb:X66363 SERINE/THREONINE-PROTEIN KINASE PCTAIRE-1 (HUMAN); mRNA sequence.

ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
N78054.1 GI:1240755  
human.  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 46)  
Hillier, L., Lennon, G., Becker, M., Donald, M.F., Chlapelli, B., Chissoe, S., Dietrich, N., Dubuque, T., Favello, A., Gish, W., Hawkins, M., Hultman, M., Kucaba, T., Lacy, M., Le, M., Le, N., Mardis, E., Moore, B., Morris, M., Parsons, J., Prange, C., Rifkin, L., Rohlfing, T., Schellenberg, K., Soares, M.B., Tan, F., Thierly-Meg, J., Trevaaskis, E., Underwood, K., Wohlmann, P., Waterston, R., Wilson, R. and Marra, M. Generation and analysis of 280,000 human expressed sequence tags  
Genome Res. 6 (9), 807-838 (1996)  
97044478  
Contact: Wilson RK  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: est@watson.wustl.edu

This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.  
Trace considered overall poor quality  
Insert Length: 1438 Std Error: 0.00  
Seq primer: reverse ET  
High quality sequence stop: 1.  
Location/Qualifiers

# FEATURES

source

1. .46  
/organism="Homo sapiens"  
/db\_xref="GDB:3797462"  
/db\_xref="taxon:9606"  
/clone="IMAGE:248216"  
/clone\_lib="Soares fetal liver spleen INFLS"  
/sex="male"  
/dev\_stage="20 week post conception fetus"  
/lab\_host="DH10B (ampicillin resistant)"  
/note="Organ: Liver and Spleen; Vector: pT73D (Pharmacia) with a modified polylinker; Site\_1: Pac I; Site\_2: Eco RI; 1st strand cDNA was primed with a Pac I - oligo(dT) primer (5' AACTGGAAGAATTAATTAAGATCTTTTTTTTTTTT 3'), double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Pac I and cloned into the Pac I and Eco RI sites of the modified pT73 vector. Library went through one round of normalization. Library constructed by Bento Soares and M. Fatima Bonaldo."

# BASE COUNT

6 a 20 c 7 g 12 t 1 others

Query Match 2.6%; Score 45; DB 14; Length 46;  
Best Local Similarity 97.8%; Pred. No. 0.78;  
Matches 45; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1700 ACTCTGCTGCTACCTGAGCCGATGTCACCTGCCACCTGTGCC 1745

DB 1 ACTCTGCTGCTACCTGAGCCGATGTCACCTGCCACCTGTGCC 46

# RESULT 2

R38968/c

LOCUS

DEFINITION R38968 28 bp mRNA linear EST 04-MAY-1995

IMAGE:25073 3' similar to gb:X66363 SERINE/THREONINE-PROTEIN KINASE

PCTAIRE-1 (HUMAN); mRNA sequence.

ACCESSION R38968

VERSION R38968.1

KEYWORDS GI:796424

SOURCE EST.

ORGANISM human.

REFERENCE Homo sapiens

AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 28)

Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M., Holman

, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M., Parsons, J.,

Rifkin, L., Rohlfing, T., Soares, M., Tan, F., Trevaskis, E., Waterston

, R., Williamson, A., Wohldmann, P. and Wilson, R.

The WashU-Merck EST Project

Unpublished (1995)

Contact: Wilson RK

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

Tel: 314 286 1800

Fax: 314 286 1810

Email: est@watson.wustl.edu

Insert Size: 1349

High quality sequence starts: 1

Source: IMAGE Consortium. LNL This clone is available royalty-free

through LNL; contact the IMAGE Consortium (info@image.llnl.gov)

for further information. Trace considered overall poor quality

Insert Length: 1349 Std Error: 0.00

Seq primer: -2lm13

High quality sequence stop: 1.

Location/Qualifiers

1. .28

FEATURES

source

/organism="Homo sapiens"  
/db\_xref="GDB:397420"  
/db\_xref="taxon:9606"  
/clone="IMAGE:25073"  
/clone\_lib="Soares infant brain INIB"  
/sex="female"  
/dev\_stage="73 days post natal"  
/lab\_host="DH10B (ampicillin resistant)"

/note="Organ: whole brain; Vector: Lfamid BA; Site\_1: Not

I; Site\_2: Hind III; 1st strand cDNA was primed with a Not

I - oligo(dT) primer (5'

AACTGGAAGAATTCGGCGCCGAGGAATTTTTTTTTTTT 3');

double-stranded cDNA was ligated to Hind III adaptors

(Pharmacia), digested with Not I and directionally cloned

into the Not I and Hind III sites of the Lfamid BA vector.

Library went through one round of normalization. Library

constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 8 a 5 c 11 g 4 t

Query Match 1.6%; Score 28; DB 14; Length 28;

Best Local Similarity 100.0%; Pred. No. 9.1e+03;

Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1716 CCGGAGCATGTTCCACCTGCCACCTGT 1743

DB 28 CCGGAGCATGTTCCACCTGCCACCTGT 1

RESULT 3

AN247978

LOCUS 2819657.5prime NIH\_MGC\_7 48 bp mRNA linear EST 07-JAN-2000

DEFINITION mRNA sequence.

ACCESSION AN247978

VERSION AN247978.1

KEYWORDS GI:6591066

SOURCE EST.

ORGANISM human.

REFERENCE Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 48)

NIH-MGC http://mgc.nci.nih.gov/.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Other ESTs: 2819657.3prime

Contact: Robert Strausberg, Ph.D.

Email: cga@bbs.rmail.nih.gov

Tissue Procurement: DCTD/DFP cDNA Library Preparation: Ling

Hong/Rubin Laboratory cDNA Library Arrayed by: The I.M.A.G.E.

Consortium (LNL) DNA Sequencing by: Berkeley MGC sequencing

Project Clone Distribution: MGC clone distribution information can

be found through the I.M.A.G.E. Consortium/LNL at:

www.bic.llnl.gov/bbrp/image/image.html Base Calling / Quality

Scores: PHRED from University of Washington Genome Center. Vector

Trimming: cross\_match from University of Washington Genome Center

PHRAP suite. Poly-T identification: patmatch.pl from Berkeley

Drosophila Genome Project. University of Washington Genome Center:

http://www.genome.washington.edu Low Quality Sequence: 8 contiguous

PHRED high quality bases following vector sequence. Very Low

Quality Sequence: Trace file contained 48 contiguous distinct peaks

following vector sequence.

Plate: L1CM2 row: C column: 18

High quality sequence stop: 8.

Location/Qualifiers

1. .48

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/clone="IMAGE:2819657"

/clone\_lib="NIH\_MGC\_7"

/tissue\_type="Small cell carcinoma"

/cell\_line="MGC3"

/lab\_host="DH10B (phage-resistant)"

# FEATURES

source







**ORCANTISM** Mus musculus  
**REFERENCE** Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.  
**AUTHORS** 1 (bases 1 to 48)  
Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,  
Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly  
,N., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.,  
and Wright,D., Weiss,R.  
**TITLE** Mouse whole genome scaffolding with paired end reads from 10kb  
plasmid inserts  
**JOURNAL** Unpublished (2000)  
**COMMENT** Contact: Robert B. Weiss  
University of Utah Genome Center  
University of Utah  
Rm 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT  
84112, USA  
Tel: 801 585 5606  
Fax: 801 585 7177  
Email: ddunn@genetics.utah.edu  
Insert Length: 10000 Std Error: 0.00  
Plate: 0036 row: F column: 16  
Seq primer: CACACAGGAACAGCTATGACC  
Class: plasmid ends  
High quality sequence stop: 48.  
Location/Qualifiers  
1  
48  
source

```

FEATURES
source
High quality sequence stop: 48.
Location/Qualifiers
1..48
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGCIM0026F16"
/clone_lib="Mouse 10kb plasmid UUGCIM library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, Tl-resistant, F-."
/note="Vector: pMD22nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pMD22 (g1147321419b1Af29072.1), a copy-number inducible derivative of plasmid p1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."
18 a 9 6 5 15

```

BASE COUNT	18 a	9 c	6 g	15 t	
ORIGIN					
Query Match		1.2%	Score 20.6;	DB 15	Length 48;
Best Local Similarity		67.4%;	Pred. NO. 7.8e+05;		
Matches 29;		Conservative 0;	Mismatches 14;	Indels 0;	Gaps 0;
QY	1499	CTACTTCATATTTCACATAAGGAGATT	CAGCTACAAAGGA	1541	
DB	5	CTAAATGCATGTTTCTACTGAATAGATGATCACCATTAAAGA	47		

RESULT 11	50 bp	linear	EST 30-AUG-2001
AU106960/c		mrna	
LOCUS			
DEFINITION	Sugano Homo sapiens cDNA library	Homo sapiens	cDNA clone
ACCESSION	AU106960		
VERSION	CAS09669		mrna sequence.
KEYWORDS	AU106960		
EST.	AU106960.1		GI:13556481
SOURCE			human





/tissue-type="moderately-differentiated adenocarcinoma"  
 /lab-host="DH10B"  
 /note="Organ: colon; Vector: PCWV-SPOFF6; Site\_1: SalI;  
 Site\_2: NotI; Cloned unidirectionally. Primer: Oligo dr.  
 Average insert size 1.7 kb. Life Technologies catalog #:  
 11531-019"

BASE COUNT 9 a 13 c 14 g 4 t  
 ORIGIN

Query Match 1.1%; Score 20; DB 9; Length 40;  
 Best Local Similarity 72.2%; Pred. No. 1e+06;  
 Matches 26; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

OY 232 GGTGGTGGTGGCGGCGTGGACCTGGAGAGAGCCCCC 267

DB 39 GGTGGTGGTGTCTTTACCAACCTGGTGGACCCCCC 4

Search completed: March 4, 2003, 00:50:32  
 Job time : 2635 secs